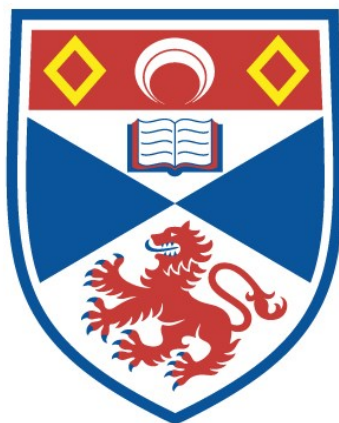


NATURALISM, NORMATIVITY, AND THE 'OPEN QUESTION' ARGUMENT

Andrew David Fisher

**A Thesis Submitted for the Degree of PhD
at the
University of St Andrews**



2003

**Full metadata for this item is available in
St Andrews Research Repository
at:**

<http://research-repository.st-andrews.ac.uk/>

Please use this identifier to cite or link to this item:

<http://hdl.handle.net/10023/13218>

This item is protected by original copyright

NATURALISM, NORMATIVITY
AND THE 'OPEN QUESTION'
ARGUMENT

by

Andrew David Fisher

A thesis submitted for the degree of

PhD

28/08/03



ProQuest Number: 10166940

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 10166940

Published by ProQuest LLC (2017). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code
Microform Edition © ProQuest LLC.

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

TH
E576

i.I, Andrew Fisher, hereby certify that this thesis, which is approximately 60,000 words in length, has been written by me, that it is the record of work carried out by me and that it has not been submitted in any previous application for a higher degree.

date 4/9/03 signature of candidate

ii.I was admitted as a research student in September 1999 and as a candidate for the degree of PhD in Philosophy; the higher study for which this is a record was carried out in the University of St Andrews between 1999 and 2003.

date 4/09/03 signature of candidate

iii.I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of PhD in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

date 4/9/03 signature of supervisor

In submitting this thesis to the University of St Andrews I understand that I am giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. I also understand that the title and abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker.

date 11/09/03... signature of candidate

.....

The University of St Andrews

Abstract

NATURALISM, NORMATIVITY
AND THE 'OPEN QUESTION'
ARGUMENT

by Andrew David Fisher

The 'open question' argument, as it has come to be known, was popularized by G. E. Moore. However, it is universally recognized that his presentation of it is unconvincing, as it is based on dubious metaphysics, semantics and epistemology. Yet, philosophers have not confined the argument to the history books, and it continues to influence and shape modern meta-ethics. This thesis asks why this is the case, and whether such an influence is justified. It focuses on three main positions, analytic naturalism, non-analytic naturalism and supernaturalism. It concludes that the 'open question' argument challenges all three.

**NATURALISM, NORMATIVITY AND THE
'OPEN QUESTION ARGUMENT'**

- (I) Moore's Open Question Argument**
- (II) Three Modern Versions of the OQA**
- (III) Moral Motivation and the OQA**
- (IV) The OQA and Rationalism**
- (V) Non-Analytic Naturalism and the OQA (I)**
- (VI) Non-Analytic Naturalism and the OQA (II)**
- (VII) Robert Adams' Supernaturalism and the OQA**
- (VIII) Dispositional Accounts of Meaning and the OQA**
- (IX) Conclusion and Suggestion.**

INTRODUCTION

G. E. Moore's *Principia Ethica* is bold. It argues that the majority of moral philosophy that predates 1903 has a problem, namely, that it rests on the false belief that 'good' is definable.¹ 'Good', Moore argues, is simple, non-natural and undefinable. His confidence in such a radical claim arises from a deceptively simple line of thinking. If we consider a definition of 'good', making it as complicated or as simple as we want, and if we ask ourselves whether the definition is correct, we should, if the definition is correct, find such questions trivial and uninteresting. However, this is not what we find. In fact, when considering the truth of such a definition we recognise that there is a very real chance that we may be wrong and that questioning the truth of our proposal is, as such, non-trivial and interesting. Moore argues that the same is true for every proposed definition of 'good' and that consequently 'good' is undefinable. This way of proceeding has come to be known as the 'open question' argument and is the focus of my thesis. Essentially, I set out to ascertain why philosophers find such an argument intuitively compelling, and whether they are right to do so.

I argue that Moore's 'open question' argument is based on dubious epistemology and metaphysics, is circular and relies on the false premise that property identity requires synonymy. This makes the topic of this thesis even more pressing. Given that such claims are typically accepted, why does the 'open question' argument continue to be so influential? I think that an answer to this is only forthcoming if we resist the urge to search for an *argument* and concede that, at best, Moore presents a *challenge* to certain accounts. Then we should ask what is the challenge? And what accounts are challenged?

¹ Moore (1993).

I start my thesis from Moore's writings. Then, after considering a number of modern versions of the 'open question' argument, I spend two chapters attempting to develop a number of open question challenges to analytic naturalism. These rely on two central ideas. First that there is a connection between moral judgment and motivation and second that there is a connection between moral judgement and practical reason.

I argue in chapters V, VI and VII that contrary to popular thinking the 'open question' argument also challenges non-analytic naturalism. There are two general challenges that I find. One starts from the fact that the naturalist relies on certain normative claims to characterise his account. The other relies on the difference in our referential intentions with moral and natural terms. As such, I conclude that Moore was indeed 'on to something' and that a moral account should respond to these open question challenges. I suggest that the correct position to adopt in light of my conclusions is what I label 'ethical nominalism'.

(I) MOORE'S OPEN QUESTION ARGUMENT

At least two things are true of George Edward Moore's *Principia Ethica*.² First, it has been unquestionably influential, changing the way philosophers study ethics.³ Second, it is undoubtedly unclear. Both of these claims will become evident in this chapter, which I split into three sections. §1 gives an overview of Moore's main aims whilst avoiding expository debates. §2 does engage, in part, in such debates when reconstructing two versions of the 'open question' argument; §3 draws some conclusions.

§1 Moore's 'Open Question' Argument.

Moore wrote that the most fundamental question in all ethics concerns how 'good' is to be defined.⁴ He claims that without an adequate answer to this question, the rest of ethics is 'as good as useless from the point of view of systematic knowledge'.⁵ He concludes that 'good' is indefinable and that, in failing to recognise this, the majority of philosophers commit the 'naturalistic fallacy'.⁶ He argues that a form of argument generally known as the 'open question' argument (OQA) demonstrates this. I develop these points below; however, before doing so, it is necessary to make some background comments concerning Moore's use of 'meaning' and 'definition'.

2 Moore (1993).

3 Darwall, Gibbard and Railton (1992) take it as the starting point for their survey of modern meta-ethics. Baldwin (1990: 66) writes: '... British ethical theory is unintelligible without reference to *PE*... [*Principia Ethica*].

4 Moore (1993: 57).

5 Moore (1993: 57).

6 As such he writes that the fallacy: '... is to be met with in almost every book on Ethics; and yet it is not recognised...' Moore (1993: 65-66).

Given Moore's view that this is the most fundamental question in all ethics, one would expect him to have a clear and distinct account of 'definition', 'meaning' and 'analysis' – he has none. His use of all three terms is vague and ambiguous.⁷ In light of this, the aim of this section is to give a 'feel' for Moore's overall project rather than attempting to connect a number of his conflicting and unrelated ideas.

Moore holds a *concept/proposition theory* of meaning. Simply put, the meaning of a predicate is the concept it stands for, and the meaning of a sentence is the proposition it denotes. However, even though this is the main theory within *Principia Ethica*, Moore sometimes writes as if he were committed to an *object* theory of meaning which holds that the meaning of an expression is the object for which it stands. Consider when he discusses what it is to define 'horse'. Here, he writes as if this requires dividing the object horse into its parts: for example, its liver, head etc.⁸ He also talks about the impossibility of dividing the object good into parts. Both of these ways of talking invite us to interpret Moore as holding that the meaning of an expression is the object it names. However, we can avoid some of this confusion concerning his account of meaning by noting that when he talks of the object referred to by 'good' or 'horse' he means object of thought, which he takes to be a concept.⁹ Once this is recognised, other passages in *Principia Ethica* become clearer. For instance when Moore writes about the

7 For example, in the opening section of *Principia Ethica*, he comments that the meaning of 'good' is the '*object or idea*' for which the word stands. Moore (1993: 58).

8 'We may mean that a certain *object*, which we all of us know, is composed in a certain manner: that it has four legs, a head, a heart, a liver etc...' Moore (1993: 60). Emphasis mine. This passage has always puzzled me as it invites questions concerning which parts are required before we are thinking about the whole horse. Moore, for instance, cites the horse's liver. In that case, do we need a horse's molecules, blood, stomach, etc? I touch on a related issue in the next section where I discuss the picture theory of ideas. However, more immediately, we can avoid some of these confusions if we note that Moore use of 'object' is ambiguous, and hence we can read the previous quotation as referring to the concept of horse. See the discussion below.

9 Although this is not explicitly in Moore (1993), we can see this view being refined in Moore (1927).

definition of the term 'chimera', reading him as claiming that its meaning is the object of thought that it denotes allows us to understand how he can talk about the terms meaning whilst acknowledging that chimera's are fictions. We can make some further remarks concerning Moore's concept account of meaning.

Moore makes it clear that concepts are not psychological. For instance, he writes:

... concepts are possible objects of thought; but that is no definition of them. It merely states that they may come into relation with a thinker; and in order that they *may* do anything, they must *be* something. It is indifferent to their nature whether anybody thinks them or not.¹⁰

And, Moore's view is that:

Apprehension of the meaning of one sentence with one meaning, differs in some respect from the apprehension of another sentence with a different meaning.¹¹

Since

Each act of apprehension is alike in respect of the fact that it is an act of apprehension, and an act of apprehension of the same kind [they can only] differ in that whereas one is

¹⁰ Cited in White (1958: 48).

¹¹ Moore (1953: 57).

the apprehension of one proposition, the other is the apprehension of a different proposition. Each proposition, therefore, can and must be distinguished both from the other proposition, and also from the act which is the apprehending of it.¹²

Thus, for Moore, the meaning of a sentence is the proposition for which it stands, and the meaning of a predicate is the concept for which it stands. It is this account of meaning which helps us understand what Moore means by 'definition' and 'analysis' and, ultimately, the OQA.

Throughout *Principia Ethica*, Moore talks about analysing 'good'. To understand this we need to ascertain what it is that he thinks we are analysing when we analyse. We can certainly ascertain that he does *not* mean. He says explicitly in *Principia Ethica* that by analysis he will not mean the consideration of verbal expressions, as this should be the concern of the lexicographer and not the philosopher.¹³ Furthermore, in various places Moore describes the analysandum as various kinds of entities, such as relations, properties, facts, and also physical objects.¹⁴ However, Alan White writes:

Above all, [Moore] speaks of the analysis of what might be grouped as 'mental', or 'psychological', entities; for in one place or another the role of *analysandum* is filled by volitions... sensations... ways of knowing... a state of desire... Where the entities to be

¹² Moore (1953:58)

¹³ Moore (1993: 58)

¹⁴ See Moore (1927: 171), (1953: 276, 265) and (1993: 59)

analysed are not, in an ordinary sense, mental or psychological, he nevertheless links them to the psychological.¹⁵

However, in *Principia Ethica* Moore most frequently talks about what is being analysed as a concept or a proposition. Moreover, given his account of meaning, when Moore writes about the analysis of 'good' he writes about analysing the meaning of 'good'. It follows then that when Moore discusses 'meaning' he is not interested in the proper usage as established by custom but:

...solely with that object or idea, which I hold, rightly or wrongly, that the word is generally used to stand for. What I want to discover is the nature of that object or idea...¹⁶

Hence, there is a further ambiguity in the notion of meaning in Moore's writing. In one reading, 'meaning' concerns proper usage as established by custom, in the other it concerns the concept or proposition for which an expression stands. However, we should read Moore in terms of the second meaning as this fits best with his account of the OQA.

According to Moore, 'good' has no definition. However, by definition he means – at least when he first introduces the topic – the process of dividing the concept of proposition that an expression names. What does this mean? Moore claims that it amounts to substituting the

¹⁵ White (1958: 51).

¹⁶ Moore (1993: 58). Emphasis added. NB. This causes further confusion when reading Moore: whereas it makes sense to say a predicate is indefinable, it makes little sense to claim that the denoted object/idea is indefinable.

concept under consideration for other concepts.¹⁷ Thus, the process of defining 'chimeara' would involve substituting the concept it denotes for other concepts: for example, the concept of a lioness, a snake, etc.

A term is then indefinable if one cannot substitute the concept it denotes for others. As such, Moore thinks that the ultimate end of all definition is a set of 'simple' concepts; concepts that are not complex and hence are ones that we cannot replace with others. He argues that good is part of such a set:

I say that [good] is not composed of any parts, which we can substitute for it in our minds when we are thinking of it. We might think just as clearly and correctly about a horse, if we thought of all its parts and their arrangement instead of thinking of the whole... but there is nothing whatsoever which we could so substitute for good; and that is what I mean, when I say that good is indefinable.¹⁸

However, Moore does not claim the *things* we call 'good' are indefinable. Hence, he writes that 'that which is good' is definable.¹⁹ It is the concept named by 'good' itself that we cannot substitute. However, it is unclear what Moore meant by a concept that cannot be substituted and unfortunately, I suggest that to gain a fuller understanding of what he did mean we need to make

¹⁷ See my discussion in the next section.

¹⁸ Moore (1993: 60).

¹⁹ Moore (1993: 61).

explicit another account of definition present in *Principia Ethica*: the *inspection* method.²⁰

Perhaps the best way to try to understand what he means is to consider an example.

Moore considers the colour yellow, which he thinks is also simple and indefinable.²¹ Essentially, he claims that due to the phenomenological qualities of yellow, one cannot substitute what we experience when we experience something as yellow in our thinking. He writes:

...light-vibrations are not themselves what we mean by yellow. *They* are not what we perceive. Indeed we should never have been able to discover their existence, unless we had first been struck by the patent difference of quality between the different colours. The most we can be entitled to say of those vibrations is that they are what corresponds in space to the yellow which we actually perceive.²²

The key to understanding this (one Moore makes explicit) is that by 'yellow' we mean what we can call a phenomenal property. It is in these terms that we can reasonably hold, as Moore does, that yellow has no parts. However, even though Moore makes this point about yellow, he does not explain what he means by a phenomenal property. He thinks good is indefinable for phenomenological reasons; via inspection of the concept denoted by 'good' we see that we cannot substitute good for any other concept. This still is far from clear and I discuss it in more detail in the next section. The main point is that there seem to be two interrelated

20 Moore does not make the distinction between division and inspection explicit.

21 He also gives the example of pleasure. Moore (1993: 64-65).

22 Moore (1993: 62).

claims made by Moore in the opening chapter of *Principia Ethica*. First, to define something we need to *divide* what it names; the second is that to define something we need to *inspect* what it names. Furthermore, he claims that 'good' is indefinable either way.

From his characterisation of simple ideas, Moore recognises a potential 'naturalistic fallacy'.²³ Assume 'A' (for instance, 'yellow') is indefinable; however, because it is always instantiated with B (for instance, light-vibrations), we mistakenly think that 'A' can be defined in terms of B.²⁴ Moore argues that ethics is particularly prone to this fallacy as it is concerned with those properties instantiated with good. For example, when investigating the good we may note that those things called 'good' are also pleasurable. This may lead to the definition of 'good' in terms of pleasure, which would be a mistake. Moore's stated aim in *Principia Ethica* is to demonstrate that most moral philosophers have committed such a fallacy.²⁵ As Moore writes:

... too many philosophers have thought that when they named those other properties [the ones instantiated with good] they were actually defining good; that these properties, in fact, were simply not 'other' but absolutely and entirely the same with goodness. This view I propose to call the 'naturalistic fallacy' ...²⁶

23 There are a number of ways of interpreting the fallacy. I discuss this below.

24 See, for example, Moore (1993: 68).

25 These philosophers include: Jeremy Bentham (pg. 70), Herbert Spencer (pg. 102), Immanuel Kant (pg. 178), John Stuart Mill (pg. 118), Stoics (pg. 165).

26 Moore (1993: 62). Put simply, the fallacy is mistakenly thinking constant conjunction is the same as identity.

Moore argues that good cannot be complex or meaningless, and therefore must be simple.²⁷ He argues that by a 'simple appeal to facts' he can demonstrate that to think otherwise is to commit a fallacy. I will now consider this argument

Moore states that 'good' could not stand for a complex idea because:

...it may be always asked, with significance, of the complex so defined, whether it is itself good.²⁸

This is Moore's first presentation of the OQA.²⁹ To elucidate these ideas Moore considers the definition of 'horse'. Consider 'horse', a noun which, as Moore claims, refers to a complex idea. He argues that we can think of 'all the parts of a horse and their arrangements *instead of* thinking of a whole.'³⁰ Imagine thinking about a horse's head on a horse's neck, covered with a horse's mane, joined to a horse's body with four horse legs, etc. Now, hold all these parts in mind. Could this be anything other than the concept of a horse? Moore thought that the answer was, obviously not; and that in such an instance, to ask whether what we are thinking about is a horse is an insignificant and 'closed' question. Moore claims that the same is true of *any* complex concept. That is, consider *all* the parts of a complex concept held 'before the mind' and it will be a 'closed' question as to whether what we are presented with is the original concept.

27 Moore (1993: 67).

28 Moore (1993: 67).

29 Moore gives a number of presentations of the OQA. I discuss these below.

30 Moore (1993: 60). Emphasis added.

Consider the definition of 'good' as that which we desire to desire, then substitute the object of your thought for this second-order desire.³¹ If 'good' does refer to a complex concept it ought to be obvious to us that to have the desire as the object of our thought is to have good as the object of our thought. However, Moore claims that it is not obvious; in his words, one can always ask with significance whether the complex (in this case the second-order desire) is itself good. This, he claims, is true for all proposed definitions of good, and therefore good cannot be a complex concept. The analyst of 'good' should:

... attentively consider with himself what is actually before his mind [in the hope that if] he will try this experiment with each suggested definition in succession, he may become expert enough to recognise that in every case he has before his mind a unique object...³²

Hence, this version of the OQA is not merely the claim that if good is not identical to some other property, say *x*, then it would be fallacious to identify good and *x*. What he attempts to do here is give a positive reason to conclude that good could not be identified with a property it is not. This is why this account differs from the version of the OQA which I label as trivially true below.

Moore claims there are only three options: 'good' denotes nothing and hence is meaningless, 'good' denotes a simple concept, or 'good' denotes a complex proposition/concept and hence is definable. Given his OQA has shown that the last option is unavailable, he concludes that 'good' must be either meaningless or denote a simple concept. However, he

³¹ Moore (1993: 67) discusses this definition.

³² Moore (1993: 68) Notice that this shifts the focus from division to inspection.

claims that it is *prima facie* implausible that 'good' is meaningless. For example, he points out that people talk about things being 'good' and are understood, and that we can communicate meaningfully with the term. Moreover, and this is where the 'inspection' account of analysis plays a role again, he claims that if we consider 'good' a concept is 'brought before our mind'. And given that the meaning of an expression is the concept/proposition it expresses, 'good' must therefore have a meaning.³³ Thus, Moore is left to conclude that 'good', must denote a simple idea. Using the OQA Moore has demonstrated that to define 'good' is to commit a fallacy.

Notice, though, that Moore has not demonstrated anything about the *ontological* status of good. The OQA demonstrates that 'good' must denote a simple object of thought, but is silent on the nature of the object. If it makes sense to talk about a simple *natural* object of thought, and Moore thinks it does, then good may be a natural object.³⁴ Hence, he needs another version of the OQA to link the discussion of the naturalistic fallacy to the ontological claims central to *Principia Ethica*, something of which Moore is apparently unaware. Consider this quotation:

In this argument the naturalistic fallacy is plainly involved. That fallacy, I explained, consists in the contention that good *means* nothing but some simple or complex notion, that can be defined in terms of *natural* qualities.³⁵

³³ See specifically, Moore (1993: 68).

³⁴ Moore claims that yellow and pleasure are both natural and simple.

³⁵ Moore (1993: 125). Emphasis added.

Evidently, Moore thought he had demonstrated in *Principia Ethica* that the naturalistic fallacy shows that good cannot be defined in terms of 'natural' properties. But he does not; at best, he only shows that 'good' denotes a simple concept.

However, he makes things even more complicated by inadvertently encouraging another reading of the naturalistic fallacy. He writes that the fallacy is committed if we identify good with something that it is not, for example, if we were to identify good with that which we desire to desire when it is in fact not *that which we desire to desire*.³⁶ I do not discuss this version, as it seems to me to be trivially true. There is obviously something wrong with an account which asserts that something is something which it is not.

In the next section I give a more detailed analyses of Moore's OQA, and I suggest that if successful it will demonstrate that 'good' is indefinable, but that it remains silent on the ontological status of good.

§2 Moore's OQA.

Moore's presentation of the OQA is confusing. Furthermore, the OQA is not an argument; it has no recognisable premises and conclusion. Thus, I suggest it is best thought of as a challenge.

Recall that the OQA is Moore's way of showing that: (I) 'good' is indefinable; (II) good is a non-natural object; (III) good cannot be anything other than what it is.³⁷ As stated above, I do not discuss (III).

³⁶ For example, see Moore (1993: 109).

³⁷ It is interesting that even though this reading is implicit and rarely discussed by Moore, it fits best with his 'epigraph' for *Principia Ethica*, that is, Bishop Butler's claim that 'Everything is what it is and not another thing'.

One might reasonably assume that (II) would be the focus of Moore's OQA. After all, his ontological position forms a central theme in *Principia Ethica*. However, Moore's aim is to demonstrate (I). Thus, even if the OQA is successful, it does not secure (II). Let us consider the argument for (I) in more detail.

Moore writes that:

It may indeed be true that what we desire to desire is always also good; perhaps, even the converse may be true: but it is very doubtful whether this is the case, and the mere fact that *we understand very well what is meant by doubting it*, shows clearly that we have two different notions before our minds.³⁸

However, one may wonder why this demonstrates anything. Why is being able to doubt a proposition about good significant? The answer lies in Moore's account of analysis. He argues that all propositions about good are synthetic.³⁹ However, why think that this is true and, more importantly, that it is relevant?

Moore argues that if someone understands a paradigmatic analytic proposition, then he or she cannot significantly question it. For example, consider: 'a bachelor is an unmarried man of marriageable age'. Given that we understand the meaning of 'bachelor', questioning the truth of the proposition is not significant.⁴⁰ However, according to Moore it is always significant and nontrivial to question a proposition about the good. In his terms: 'we understand very well what

38 Moore (1993: 68). Emphasis added.

39 Moore (1993: 58).

40 Of course, modern philosophy of language has shown that we can. For example, the question 'is the Pope a bachelor?' appears 'open'. Nevertheless, in this chapter - *pace* Quine - I assume that the analytic/synthetic distinction is meaningful.

is meant by doubting it'.⁴¹ He concludes that all proposed propositions about the good must be synthetic.⁴²

However, as Moore later recognised, to accept this line of reasoning means his position leads to a 'paradox of analysis'.⁴³ If he thinks that because one can doubt a proposition the conceptual analysis must be incorrect, he must also think that all correct conceptual analyses are obvious and hence uninformative. A way of trying to understand why Moore is committed to the paradoxical conclusion follows. Moore claims that to define a term was to substitute the object of thought it denotes with its components. However, to do this we must *already* know which parts form the whole. I glossed over this point when I outlined Moore's account of the definition of 'horse'. In such an analysis, we substitute the concept of horse for a horse's head, a horse's mane, a horse's hooves, etc. However, for this 'definition' to be successful we must know two things: first, what a *horse's* head, a *horse's* mane, *horse's* hooves are, and so on. Second, we must know when we have completed the substitution.⁴⁴ Arguably, then to give a satisfactory definition of 'horse' we must have complete propositional knowledge concerning the meaning of 'horse'. For it is only then that we could both know the parts we have are parts of the whole, and that they form a complete concept. Given the nature of propositional knowledge, Moore is thus committed to the conclusion that all conceptual analysis must be obvious and

41 Moore (1993: 68).

42 Note though that at best this only invites the naturalist to provide a proposition about good that *is* analytic. That is, it does not establish that we could *never* find a proposition about good that we could not doubt. Thus, as stated this version of the OQA is best thought of as a challenge, rather than an argument.

43 See Moore (1942).

44 We could put this in terms of his 'inspection' account of definition: for us to 'see' the concept more clearly, we have to already know what we are meant to be looking at, and for.

uninformative. However, consider for example, Frege's analysis of the concept of number.⁴⁵ This is an unobvious and informative conceptual analysis; thus, the paradox of analysis appears to be a *reductio* of Moore's account of analysis.

A position concerning conceptual analysis could avoid the paradox if it recognised the difference between 'knowledge how' and 'knowledge that'. A simple example of the difference is that even though I know how to ride my bicycle, I have no propositional knowledge that represents this ability. I could not, for instance, describe how the centripetal forces react to keep me balanced as I swing around a hairpin bend. In general terms, though, analysing 'A' requires grasping the meaning of 'A'. This need not entail having complete propositional knowledge concerning 'A'. Grasping its meaning may include knowing how to do various things with 'A'. Hence, it could be a significant question to ask whether 'A is B', even if we know the meaning of 'A', because we may not know *that* 'A is B'. Therefore, it does not follow that from a proposition being analytic we could fail to know what it means to doubt it. So, by avoiding the paradox of analysis, Moore's line of reasoning fails to justify his claim that all propositions about good are synthetic.

Consider Moore's reasoning again. He argues that one can always ask significant and non-trivial questions concerning propositions about the good; hence, they all are synthetic. He also argues that property identity requires analyticity. So, if all propositions about good are synthetic, we cannot reduce the property good to another property. Now, the 'paradox of analysis' already shows the OQA cannot sustain that conclusion. But an important further point

⁴⁵ It could be questioned whether what Frege provides is an analysis; however, most philosophers would count this as an analysis, and so do I. Moore would, then, have to demonstrate why he, holding the less popular position, was correct to deny that it was an analysis.

is that philosophers now generally accept that property identity does not require *a priori* meaning equivalence.⁴⁶ Thus, even if Moore's claim about analysis were correct, this version of the OQA would not establish that good is a simple non-natural property.

Even given the failure of this version of the OQA it has been hugely influential. For instance, Darwall, Gibbard and Railton write:

However readily we now reject as antiquated his views in semantics and epistemology, it seems impossible to deny that Moore was on to something.⁴⁷

Their suggestion is that it is 'impossible to deny that Moore was onto something' because people whom we judge to be psychologically and linguistically competent do find questions concerning naturalistic accounts open. Yet, if the best account of Moore's OQA is as DGR state, then the philosopher attempting to define 'good' would most likely not be worried. After all, people are often confused, and so there is no reason to think that people's linguistic intuitions reflect the truth.⁴⁸ Hence, for Moore's OQA to be a genuine challenge we require an account to justify the claim that open question intuitions have evidential weight. I find two accounts in *Principia Ethica*.⁴⁹

The first uses Moore's phenomenological claims. However, I judge that even though these are more plausible than they first appear, they still fail to justify the OQA as they require a

46 See, for example, Putnam (1975) and Kripke (1980); I discuss this in chapters V and VI.

47 Darwall, Gibbard and Railton (1992: 3).

48 Perhaps, though the naturalist cannot dismiss this approach so easily. I discuss this in the next three chapters.

49 However, note that Moore's OQA has now been reduced from an argument demonstrating that good must be a simple non-natural property, to a challenge against the proposal that 'good' can be defined.

coherent foundation in epistemology and connection to metaphysics that is absent in *Principia Ethica*.

Moore writes:

And if [one] will try this experiment [questioning a proposed definition of 'good'] with each suggested definition in succession, he may become expert enough to recognise that in every case he has *before his mind a unique object*...⁵⁰

And,

Whenever [one] thinks of 'intrinsic value,' or 'intrinsic worth,' or says that a thing 'ought to exist,' he has *before his mind the unique object* – the unique property of things – which I mean by 'good.'⁵¹

As stated, this suggests a phenomenological reading of the OQA. However, can we make sense of the claim that to think about good is to have a 'unique object before our mind'? I argue below that we can if we ascribe to Moore a specific account of consciousness, an account which also lends plausibility to his intuitionism.

One way of attempting to understand what it is to have a 'unique object before our mind' is to consider Moore's conception of analysis. The most frequent account of analysis in

⁵⁰ Moore (1993: 68). Emphasis added.

⁵¹ Moore (1993: 68). Emphasis added. I discuss below the obvious worry that this *is* a definition of 'good'.

Principia Ethica claims that to analyse a term is to inspect the concept it stands for, and divide it into its parts. However, what this means is also unclear; perhaps a way of understanding it is as follows. To analyse a term, we bring before our mind the concept denoted by that term and we 'concentrate' on it. We 'focus' our attention on it to try and 'see' it more clearly, and in so doing its composition is revealed to us.⁵² How should we make sense of this? Well, possibly this way of talking makes sense if we adopt a theory which holds that a concept is a *picture*: the concept of a horse is a picture of a horse, the concept of a house is a picture of a house, etc.

This seems to be a useful way of proceeding as, for example, we know what it is for something to be part of a picture that represents a horse. We may be presented with a photo of a friend's horse, and we could circle the mane, cover up everything except the head, cut out the torso, etc. We could physically isolate parts of the picture of the horse. When we are asked what parts make up a picture of a horse, we do not frown and shrug our shoulders: we know how to answer.⁵³

Perhaps then if concepts are pictures we can apply the same reasoning to them. A concept is a concept of a horse because it pictures one, and if we concentrate on this concept, we can understand what a part of it would be. In effect, in our thoughts we can go through the same process as we would with the real picture. For example, we might bring the concept before our mind, and think of part of the picture of the horse disappearing, leaving say the head. Now, underlying the plausibility of this claim is the thought that just as pictures in the external world are made of colours, etc, so are our concepts, and that, as a result of this, we can inspect our

⁵² I am well aware that this language is highly metaphorical, but there seems to be no other way to account for Moore's discussion of analysis by introspection; we will always *have* to talk in terms of 'seeing' of the 'mind's eye' etc.

⁵³ For instance, we could describe to someone the parts of a picture of a horse, in order that he/she could draw it.

concepts and also substitute the concepts for their components. Hence, I suggest that if we think of concepts as pictures, we can start to understand what Moore means by his inspection and division account of analysis. The suggestion, then, is that Moore's phenomenological version of the OQA requires us to conceive of concepts as images in the mind.

Consider how this version of the OQA would run. In contrast to the concepts of houses and horses, it seems inappropriate to say that the concepts denoted by 'good', 'yellow' and 'pleasure' are pictures of good, yellow and pleasure. This way of talking makes no sense - just consider being asked what parts make up 'yellow' or 'good'; in this case, we would probably frown, shrug our shoulders, and be unable to answer. Hence, Moore would argue that when we bring good, yellow and pleasure before our mind we are presented with things that are unique, are simple, and indefinable. That is, we cannot inspect them in more detail to ascertain their parts and we cannot attempt to divide them into parts. Hence, although they are concepts, they are different in type from other concepts, such as the concept of a horse, a house, or a chimera as they are not pictures. This then helps us start to comprehend Moore's phenomenological version of the OQA.

However, there is one major problem with this account of the nature of concepts - it is false. In fact, I suggest that to think that the meaning of a term is the concept that it denotes, and that what a concept is is a picture representing something, is so clearly false that people have concluded that Moore's phenomenological OQA is irrelevant and hopeless.⁵⁴

⁵⁴ Hence, typically articles on the OQA focus on meta-linguistic issues; for example, Ball (1991) and Kalderon (forthcoming).

To summarize, I suggested that one natural way to understand Moore's notion of analysis by division and inspection is to adopt an account of concepts as picture. I claimed that people do implicitly and explicitly ascribe such an account to Moore. And that consequently, because philosophers correctly see this view of the nature of concepts hopeless, they reject Moore's phenomenological version of the OQA. However, in what follows, I show that Moore would not have accepted such an interpretation of his position.

It is interesting and informative to note that Moore states in the preface to the first edition of *Principia Ethica* that within Franz Brentano's work he finds:

... opinions far more closely resembling my own, than those of any other ethical writer with whom I am acquainted.⁵⁵

Arguably, the most influential of Brentano's opinions concerns intentions. Consider the following quotation from him:

Every psychological phenomenon is characterised by...intentional inherent existence of... an object (*by which we do not mean a reality*)... In the idea something is conceived, in the judgement something is recognised or discovered, in loving loved, in hating hated, in desiring desired, and so on.⁵⁶

⁵⁵ Moore (1993: 36). He cites specifically Brentano's (1969) *Origin of the Knowledge of Right and Wrong*.

⁵⁶ Brentano (1973: 88). Emphasis added. A similar quotation can be found in Brentano (1969: 14).

Brentano's basic thought, then, is that all thinking is thinking about a real, or unreal, object. I will call these objects *intentional objects*.⁵⁷ I suggest that for Moore also, to think about a horse is to think about an intentional object, to think about a chimera is to think about an intentional object, and importantly, to think about good is to think about an intentional object.

Of course, we could read this suggestion as neither clarifying nor accounting for anything. How does knowing this advance our understanding of Moorean 'analysis', 'definition' and the OQA? That is, discussion of intentional objects is compatible with the previous unhelpful way of reading Moore's phenomenological claims. For example, thinking about a horse may be both thinking about an intentional object and thinking about a *picture* precisely because the intentional object *is* a picture. However, I suggest below that Moore does not think about intentional objects in this way. Consequently, I argue that there is more reason to accept his phenomenological claims than this first reading suggests.

In the same year that he published *Principia Ethica*, Moore also published the influential paper 'The Refutation of Idealism'.⁵⁸ I claim that ideas found in the latter can shed light on the former.⁵⁹ What interests me in 'The Refutation of Idealism' is Moore's account of consciousness,⁶⁰ in which he claims that introspection never 'presents us' with consciousness. He argues that this is because consciousness is not a *thing*. This helps us make sense of the role of intentional objects in the phenomenological version of the OQA, or so I will argue. Consider Moore's remarks on the nature of consciousness:

⁵⁷ Hence, intentional objects can be real or unreal.

⁵⁸ Moore (1903).

⁵⁹ Baldwin (1990) certainly thinks this is true.

⁶⁰ A similar account can be found in Brentano (1973) and Sartre (1956).

... the moment we try to fix our attention upon consciousness and to see *what*, distinctly, it is, it seems to vanish: it seems as if we had before us a mere emptiness. When we try to introspect the sensation of blue, all we can see is the blue: the other element is as if it were *diaphanous*.⁶¹

It follows that if consciousness is not a thing then nothing could be *in* it. Specifically, it could not contain ideas conceived as pictures. Hence:

What I wish to point out is (1) that we have no reason for supposing that there are such things as mental images at all...⁶²

And:

To have in your mind 'knowledge' of blue, is *not* to have in your mind a 'thing' or 'image' of which blue is the content. To be aware of the sensation of blue is *not* to be aware of a mental image – of a 'thing,' of which 'blue' and some other element are constituent parts in the same sense in which blue and glass are constituents of a blue bead.⁶³

61 Moore (1903: 450). Emphasis added.

62 Moore (1903: 449).

63 Moore (1903: 449).

Moore's claims that there is no question of how we are to 'get outside the circle of our own ideas and sensations as:

Merely to have a sensation is already to *be* outside that circle. It is to know something which is as truly and really *not* a part of *my* experience, as anything which I can ever know.⁶⁴

I think then that we can reasonably reject the claim that Moore's phenomenological account of the OQA demands construing concepts as images, and as such we cannot dismiss his OQA so easily.⁶⁵ In fact it turns out that not only does ascribing to Moore (1993) this account of consciousness allows us to reject naïve accounts of Moore's OQA, but we can also respond to another type of worry raised by Hume:

Take any action allow'd to be vicious: Wilful murder, for instance. Examine it in all lights, and see if you can find that matter of fact, or real existence, which you call *vice*. In which-ever way you take it, you find only certain passions, motives, volitions and thoughts. There is no other matter of fact in the case. The vice entirely escapes you, as long as you consider the object.⁶⁶

⁶⁴ Moore (1903: 451).

⁶⁵ In fact this view rules out any account where there is an item in consciousness, for example, a mental 'language'.

⁶⁶ Hume (1978: 468).

This worry does seem to hold for Moore if we regard consciousness as limited to sensations and introspection, for then it seems to follow that good is not an object of consciousness, and Moore's position has a problem.⁶⁷ However, given my suggestion concerning Moore's account of consciousness Moore need not be forced into accepting this conclusion. As Panayot Butchvarov states concerning Moore's account of consciousness:

An immediate consequence of this conception is that any state of consciousness owes its distinctive character solely to its object, or more precisely to the object as it is given to that state of consciousness. The consciousness in itself has no character, no content, is pure transparence, translucence; if considered in abstraction from its object it is *nothing*...⁶⁸

If, then, there is no intrinsic qualitative nature of consciousness – it is not a thing - any state of consciousness owes its distinctive character to its intentional objects.⁶⁹ Given that Moore holds this account of consciousness, as 'The Refutation of Idealism' seems to suggest, we can say that to be in a state of ethical consciousness means only to be consciousness of the ethical or moral objects or properties, such as good. And it cannot consists in the occurrence of:

67 This would then cause a problem for his intuitionism.

68 Butchvarov (1982: 54).

69 See my discussion of conscience and qualia in chapter IX.

...what Ryle would have described as ghostly events within oneself, having some intrinsic character, presumably “emotional” or “attitudinal,” and only causally related, if at all, to their so-called objects.⁷⁰

To morally approve of something then can only amount to being conscious of it as morally good, to be conscious of its moral goodness, even if this were illusory. Ethical consciousness must be consciousness of certain objective values, of certain properties of its objects, even if these were in some sense unreal, since there is nothing else ethical consciousness could be. Moreover, good in this sense must be objective in that it could not be *in* consciousness. Importantly this account follows from the nature of conscience which I think fits best with what Moore was writing at the time of *Principia Ethica*, and *not* from any specific account of ethical or moral consciousness.

My claims then are that one must read *Principia Ethica* as holding a doctrine of intentional objects, and as rejecting the claim that these intentional objects can be *in consciousness*.⁷¹ It follows that Moore is not committed to some ‘ghostly realm’ in which we experience good. If we think about good, then we are thinking about an intentional object. This in turn makes his phenomenological claim more plausible than it may seem at first.

The account of good as an intentional object is therefore promising. However, although these remarks help, they still leave a problem. For, if the OQA is to challenge naturalism, Moore must answer the metaphysical question concerning what sort of property good is. Of course, he

⁷⁰ Butchvarov (1982: 55).

⁷¹ This seems correct. For an excellent discussion of issues surrounding consciousness and indirect realism see McCulloch (2003).

claims that good is simple, non-natural, and indefinable, but this merely returns us to the original problems concerning what it is for something to be non-natural, simple and indefinable. In fact, Moore himself claims that throughout his career his distinction between natural/non-natural, definable/indefinable, and simple/complex which are now so vital, remained unclear.⁷² Hence, it is improbable that Moore's *Principia Ethica* can be read as giving a coherent and adequate metaphysical account of good.⁷³

In conclusion, if Moore wants to justify his phenomenological claims he must adopt the account of consciousness and intentional objects discussed. However, as he is committed to the claim that any state of consciousness owes its distinctive character solely to its object, or more precisely to the object as it is given to that state of consciousness, he needs to present a clear and coherent metaphysical and epistemological position, something that is not present in *Principia Ethica*. I leave it that if Moore gave such a foundation he would have a challenge to the naturalist, one that does not depend on the implausible phenomenological claims I discussed.

My second suggested reading of Moore's OQA develops from a quotation:⁷⁴

Whenever [one] thinks of 'intrinsic value,' or 'intrinsic worth,' or says that a thing 'ought to exist,' he has *before his mind* the unique object – the unique property of things – which I mean by 'good.'⁷⁵

72 Moore (1942).

73 We could of course reconstruct an account; it is just that in doing so we will distort some part of *Principia Ethica*.

74 Baldwin (1990) and White (1958) also make this suggestion.

75 Moore (1993: 68). Emphasis added.

On the surface, this is an odd thing for Moore to write. He claims that the meaning of a term is whatever it denotes.⁷⁶ From the quotation then, it follows that 'good', 'intrinsic worth', 'intrinsic value' and 'ought to exist' have the same meaning. This would mean that 'good' is definable, something which Moore frequently reminds us is impossible. However, putting this issue to one side, the quotation is useful as it suggests another way of conceiving of the OQA.

There is an intuitive link between the terms cited in the quotation. If someone judges that a state of affairs is *good*, that it has either *intrinsic value* or *intrinsic worth*, then arguably 'this issues in a specification of what, in a particular situation, one *ought* to do'.⁷⁷ Now, if this is correct then this may account for people's open question intuitions. Arguably, if we consider a naturalistic state of affairs, say the amount of pleasure that a certain action would give, this does not specify what we ought to do. If we consider good, and we consider whether something would give us pleasure, the former 'tells us' what we ought to do, whereas the latter does not. Put very crudely, natural facts are 'cold' and 'mute', whereas good is dynamic and 'guides us' in our actions. Hence, because our intuitions reflect this gap, we resist any identification of good with natural properties – such a definition would be open. What this means is that if an analytic naturalist wanted to resist this version of the OQA, he would have to demonstrate that his proposed reduction does in fact tell us what we ought to do..

There is more textual evidence to support this reading of the OQA. For example Moore writes:

⁷⁶ See for example, (1903: 68).

⁷⁷ Baldwin (1990: 76). Emphasis added. Notice that talk of a state of affairs as good seems to conflict with the claim that good is an attributive adjective. I discuss this in the next chapter.

The whole object of Mill's book is to help us to discover what *we ought to do*; but, in fact, by attempting to define the meaning of this 'ought,' he has completely debarred himself from ever fulfilling that object: he has confined himself to telling us what *we do* do.⁷⁸

Moore's point is that John Stuart Mill's account fails because it must always reduce the prescriptive 'ought' to descriptive claims. Moreover, this means Mill's proposed account of 'good' is missing something, and hence questions concerning it will always be open. Specifically, Mill's account can only ever tell us what we *do* do, and can never tell us what we *ought* to do.

However, for this reading of the OQA to be correct Moore must hold that no naturalistic judgements are 'ought-implying'.⁷⁹ Yet such a reading fits uneasily with other claims he makes concerning the reducibility of 'ought' judgements to naturalistic judgements. For instance, Moore states that there is 'at least one "ought-implying" natural property', and in his paper 'The Conception of Intrinsic Value' he addresses some of the problems arising from adopting this relation between 'is' and 'ought'.⁸⁰ ⁸¹Thus, this reading of Moore's OQA seems problematic. We either have to reject some of Moore's claims concerning the reducibility of 'ought'

78 Moore (1993: 125). This quotation is cited by Baldwin (1990: 89). Emphasis mine.

79 That is, to justify his claim that there is a difference between 'telling us what we ought to do' and 'telling us what we *do* do.'

80 Moore (1942: 604).

81 Moore (1922a) and Moore (1942).

judgements to 'is' judgements, or reject the claim that his OQA is based on the irreducibility of 'ought' judgements.⁸²

§3 Conclusion

Moore's OQA has been hugely influential, generating a vast body of literature. However, Moore's presentation of it is neither clear nor convincing. For his argument to be successful, one or more of these must be true: conceptual analysis is obvious and uninformative. We have a unique property before our mind when thinking about good; there is an is/ought gap. The first claim leads to a 'paradox of analysis'. The second, although less naïve than a superficial reading reveals, depends on an adequate epistemological and metaphysical account of good, something which Moore does not provide. The third brings the OQA into conflict with the rest of Moore's moral philosophy.

However, even if Moore's OQA is successful, the conclusion it generates highlights an ambiguity in Moore's presentation of 'the naturalistic fallacy'. The OQA cannot demonstrate that good is a non-natural property. At best, it challenges the analytic naturalist's claim that good is definable. As such, the OQA, even if successful, becomes redundant in establishing the overall ontological position in *Principia Ethica*. However, even given Moore's failure, I will argue that the OQA is a genuine argumentative device which puts the burden of proof on the analytic naturalist.

⁸² For an excellent and detailed discussion of this see Baldwin (1990: 89-90)

(II) THREE MODERN VERSIONS OF THE OQA

This chapter deals with three recent accounts of the OQA. These have a number of things in common. First, all avoid giving an exposition of Moore's work and concede that in its original form the OQA is unsuccessful. Second, all claim that the OQA is an argumentative device rather than a 'knock down' argument. Third, all think that the OQA challenges *only* analytic naturalism. I do not engage in a critique of these accounts; I outline them because they inform the discussion in the following chapters.⁸³

§1. Stephen Ball's 'Linguistic Intuitions'.

Recently, Stephen Ball developed an account of the OQA based on pre-philosophic intuitions about moral language.⁸⁴ Stating the intention of his paper he writes:

... it will be contended that a linguistic version of Moore's argument, when appropriately developed, provides at least useful *evidence* against naturalistic theories...⁸⁵

Ball sets out the OQA as follows:

- (1) Any given naturalistic definition gives rise to an open question.
- (2) Therefore, no such definition can be *analytically* (or definitionally) true.

⁸³ Ball focuses on meta-linguistics, Baldwin on how we engage with the world, DGR on the result of engagement.

⁸⁴ Ball (1991).

⁸⁵ Ball (1991: 4).

- (3) Therefore, no moral term refers to the *same property* as a natural term.⁸⁶

I do not discuss the move from (2) to (3). The focus of this section is (1) and (2).

Traditionally, philosophers have rejected the move from (1) to (2) as circular or invalid,⁸⁷ circular because (1) merely assumes the truth of (2). Denying that a proposition is analytic is merely a sophisticated way of stating that questions concerning it are open.⁸⁸ The move is considered invalid because the criterion for an 'open question' referred to in (1) is the fallible intuitions of native speakers of English.⁸⁹ I consider these worries in turn.

Circularity: Ball suggests that pre-philosophic intuitions can provide a non-circular test of analyticity. Imagine attempting to demonstrate the idea that 'good is pleasure' is synthetic. We ask people about the proposition and we find that typically they feel that such questions are nontrivial and significant. Hence, we conclude that the proposition is not analytic. Ball claims that if it is non-philosophers that are asked then this way of proceeding is not circular. This is because this 'data' – the responses of the people asked – does not presuppose any meta-linguistic theorizing. If this is the case, the truth of (1) – the fact that people respond in the way that they do – is not based on the presumption of the truth of (2). How could it? What do the pre-philosophical know about analyticity? Ball claims that making (1) into a 'test' in this way means that the OQA avoids the standard circularity objection.⁹⁰

⁸⁶ Ball (1991: 6).

⁸⁷ For example, see Frankena (1939) and Brandt (1979).

⁸⁸ This is essentially the main point made by Frankena (1939).

⁸⁹ Ball (1991: 7). This also means that stating (1) is controversial, for it depends on the ability to infer from the subjective linguistic intuitions of native speakers.

⁹⁰ For instance, Ball writes: '... Moore's linguistic *experiment*, in (1), is to be thought of as conducted in the laboratory of the real world with ordinary native speakers, and the philosophic claim of Moore's type of *argument*, in the move from (1) to (2), is that the "raw data" so collected, i.e., what might be called the "raw feels" of linguistic experience, are relevant for

Another consideration recommends Ball's way of proceeding. *Asking* native speakers certain questions can provide more information than merely observing how they behave. Ball gives an example from Charles Stevenson.⁹¹ Stevenson argues that we need not conclude that universalisability is analytically part of the meaning of moral terms from the fact that people do not ordinarily make non-universalisable moral statements. Perhaps, it is just that such statements are impractical, or pointless:

There is reason to believe, then, that checking speakers' attitudes, or feelings of reluctance about making such denials – which are otherwise synthetically false or psychologically impractical – is at least an important kind of data which is *relevant evidence* for evaluating naturalistic claims of analyticity.⁹²

Given these points it seems that the circularity objection fails. It is not circular to move from (1) to (2).

Invalidity: These claims have direct relevance to the charge that the OQA is invalid. Ball is not proposing that one can *deduce* that certain propositions are not analytic from the 'data' concerning native speaker's intuitive feelings. The pre-philosophic intuitions are merely *inductive* evidence. Hence, if valid means *deductively* valid then this version of the OQA avoids the 'invalidity' criticism. However, there is another reading of 'invalid'. Because people's pre-

establishing a conclusion about meaning analyticity.' Ball (1991:17). Notice though that we can take 'on board' the main claim that the truth of (1) does not presuppose the truth of (2) without adopting all Ball states in this quotation. He also takes his view as an extension of the Quinean view that the analytic/synthetic distinction rests ultimately on behaviour criterion as to how statements intuitively *feel*. Specifically, statements are analytic to the degree that there is a reluctance to revise them.

See Quine (1951).

⁹¹ Ball (1991: 20).

⁹² Ball (1991: 20).

philosophic intuitions are possibly wrong, one cannot draw any conclusions from them. To do so would be 'invalid' in this revised sense. Recall that although the above way of proceeding rules out one version of invalidity, it leaves a challenge. Unless it can be shown that pre-philosophic intuitions are reliable, then it would be bad induction to base any conclusion on them.

Ball gives two reasons to distrust people's intuitions. First, people could be systematically misled in their intuitions, even on reflection. For example, perhaps people still have outdated religious views that reify moral properties as supernatural. This then creates the impression that moral terms do not mean the same as natural terms. Second, the complexity of proposed reductions may confuse people such that they find questions concerning them open. For instance, people's ignorance about Frank Ramsay's account of reduction employed by Frank Jackson would presumably lead them to question whether such an account is correct.⁹³ However, we do not think that as such Jackson's account fails to be true. Both these considerations demonstrate that one cannot simply make unqualified inference based on people's intuitions. Hence, for Ball's OQA to be convincing he has to dismiss such worries.

Ball argues that even if people's ordinary linguistic intuitions are confused, this does not challenge his account. First, most of the proposed 'distortions' are not genuine, by which he means that they are things which the naturalist attempts to capture in their reduction. For example, concerning the confused notion of objectivity as indicated by Mackie, he writes:

⁹³ Jackson (1998).

Ethical naturalism, too, attaches at least a *kind* of “objective,” factual meaning to moral terms, and thus native speakers need not be misled by the test to reject the latter merely due to a feeling that it leaves out this component.⁹⁴

Therefore, even if there are distortions, the set of intuitions that have been corrected includes more than one may think. Hence, typically it is fine to make inductive inferences about analyticity from the set of intuitions.

Second, Ball claims that even the genuine distortions will dissipate after rational reflection. Hence, the ‘data’ from which we draw inference^{e?} about analyticity should only include the responses of native speakers who have reflected rationally on the proposed reduction. This is because after reflection the data will not include, or would be unlikely to include, distorted intuitions.⁹⁵ He writes:

On the other hand, if Mackie’s metaphysical objectivity is, improperly, part of the “meaning” of ordinary moral terms, this conceptual mirage may be dissipated by rational “reflection”...⁹⁶

To the charge of complexity, Ball has three responses. First, he suggests that even if people were aware of the complexity of a proposed reduction, they may still find questions

94 Ball (1991:26).

95 Of course, he has said that the distortions could remain even after reflection; hence, he cannot claim that after reflection *all* intuitions will lack distortion.

96 So, the thought continues, the open-question is free to operate unambiguously as a test of the residual meaning proffered by a naturalistic definition. Ball (1991:26).

open. That is, the complexity is only one factor that helps account for peoples' linguistic intuitions, and perhaps it is not even part of the open question test itself. The open question test only calls on speakers to be able to detail their feelings about 'openness'.⁹⁷

Second, Ball reaffirms that even if two expressions are synonymous but differ in complexity, the open question test can still have relevance *if* synonymous expressions tend to be similarly complex. Moreover, he claims most naturalistic accounts do show this to be true. For example, 'good is pleasure' or 'good is what one desires to desire'.

Third, Ball writes:

... though ordinary people may well be surprised about the needed complexity in a proper philosophic analysis of moral language, there is no reason why this complexity should not be incorporated into their linguistic intuitions, even if they cannot articulate, or are wrong about, the details.⁹⁸

The point is simply that people's intuitions can reflect knowledge of the meaning of terms without any accompanying propositional knowledge about the meaning. Ball claims that presumably, then, the main problem with many naturalist definitions is that they are in fact too simple. He gives two reasons why this might be so. First, moral language has a multi-factual dimension including desires, social approval, etc., whereas the naturalist defines moral terms with a single factual dimension of this kind. Second, moral terms have a prescriptive force that

⁹⁷ Ball (1991:26).

⁹⁸ Ball (1991: 27). This looks an odd thing to say. For if people may 'be surprised' then it is unlikely that it will be incorporated into their pre-philosophical intuitions.

is not contained in simple, uni-dimensional naturalistic definitions.⁹⁹ These suggestions are developed in sections two and three.

§2. Thomas Baldwin's Suggestion.

Thomas Baldwin has two suggestions concerning versions of the OQA that also start from our semantic intuitions, and as such 'sidestep' issues relating to the 'paradox of analysis'.¹⁰⁰

Central to both suggestions is the idea that the starting point for the OQA ought to be the phenomenological fact that people do find questions concerning reductions significant. Baldwin writes:

I think... that we can to some extent sidestep these issues by assuming, on Moore's behalf, that if a conceptual analysis is correct, then, once we have encountered it, it should come to seem to us entirely appropriate to guide our thoughts and judgements in accordance with it, even if at first the analysis strikes us as unobvious; and Moore's objection to proposed analyses of intrinsic value is precisely that we do not find ourselves able to move to this reflective assimilation of them.¹⁰¹

Most likely the naturalist would argue that the open question 'phenomenon' is an illusion - a distortion, perhaps of religious belief. The reason that people are not happy to let the analysis

⁹⁹ Ball (1991: 27).

¹⁰⁰ Baldwin (1993).

¹⁰¹ Baldwin (1993a: xix).

'guide their thoughts and actions' is not because the analysis is incorrect; rather, they are systematically confused about the criterion for a correct reduction. Consequently, if Baldwin could give an account of ethical concepts that seeks to establish why reductive analyses are incorrect, then the OQA would be a genuine challenge. Before considering how he proceeds, we must clarify the methodology.

Baldwin's suggestion starts from a problem he recognises in Moore's work. In *Principia Ethica*, when discussing his ideal utilitarianism, Moore writes about states of affairs being good.¹⁰² Typically we use the concept 'good' as attributive, as in 'good knife' or 'good bike', where the noun ('knife', 'bike') which follows 'good' specifies what the thing said to be good is good *as*, and thereby indicates the standards by reference to which it has been evaluated. However, for Moore, goodness is a simple non-natural property that cannot readily accommodate this feature of the concept. So, there seems to be an immediate problem in understanding Moore's account of good. In fact, a number of philosophers have rejected Moore's ethics precisely because of this type of worry.¹⁰³ However, Baldwin says that, if properly understood, Moore's account could accommodate this feature of the concept. Central to Baldwin's idea is the recognition that it is the goodness of states of affairs that is the main focus of Moore's ethical theory.¹⁰⁴ How is Baldwin going to show that this focus is consistent with the fact that in its typical use the concept good is attributive? He starts with this claim:

¹⁰² Moore (1903: 235-23).

¹⁰³ Most famously, Geach (1956).

¹⁰⁴ Moore (1993: 239).

... if a phrase 'conveys a standard of goodness' as long as there is a conventional way in which such a standard is identified by the use of the phrase, then phrases such as 'event' and 'state of affairs' do convey a standard of goodness.¹⁰⁵

Baldwin considers the familiar phrase 'good weather'. We readily understand this phrase without investigating the context of utterance, even though good weather is not good *as weather*. Good weather is simply weather that is good for people affected by it, and this is weather which suits their interests.¹⁰⁶ It is only when there are no such obvious implications for human interests that we have to return to the context of utterance for further clues in how to understand what is being said.¹⁰⁷ Hence, if we do not use 'good' in a phrase that conveys a standard of goodness by the simple principle that a good *X* is something that is good *as an X*, then a standard for goodness is:

... provided by reference to the implications for the *interests* of those affected by, or concerned with, that of which goodness is predicated. This suggestion can be readily applied to the interpretation of 'good states of affairs', and since it yields a standard of goodness which the phrase conventionally conveys... [the criticism considered above concerning how we can reconcile Moore's account with the fact that in its typical use the concept good is attributive] is avoided.¹⁰⁸

¹⁰⁵ Baldwin (1990: 73).

¹⁰⁶ And: '... where interests conflict, we relativise our judgements, distinguishing between, say, weather that is good for farmers and whether that is good for holiday-makers.' Baldwin (1990: 73)

¹⁰⁷ For example, when someone says 'that is a good cloud'. See Baldwin (1990: 73-74).

¹⁰⁸ Baldwin (1990: 74). Emphasis added.

Even if this way of proceeding saves Moore's account, there is a fundamental problem in ascribing such a view to Moore. The thought was that we can combine a recognition of the attributiveness of 'good' with the focus on the goodness of states of affairs required by Moore's ideal utilitarianism by assigning a central role to human interests in the evaluative states of affairs. However, Moore:

... will reject any such role for the concept of human interest, on the grounds that it gives rise to an unacceptable form of ethical naturalism, and it is this rejection, rather than his simple failure to recognise that attributiveness of 'good', which makes his conception of goodness so problematic.¹⁰⁹

Why would Moore find it incompatible with his account? Well, because for Moore there is no distinction between something that is good *for* an agent, and something that is good in itself. Just consider the problem that Moore finds in Henry Sidgwick's egoism:

In short it is plain that the addition of 'for him,' 'for me' to such words as 'ultimate rational end,' 'good,' 'important' *can introduce nothing but confusion*. The only possible

¹⁰⁹ Baldwin (1990: 74).

reason that can justify any action is that by it the greatest possible amount of what is *good absolutely* should be realised.¹¹⁰

For Moore, 'good for me' can *only* be understood as applying to states which are good in themselves and somehow concern me.¹¹¹ However, as Baldwin notes what this means is that Moore cannot accommodate the attributive nature of 'good' in the way he suggests. As the evaluation of states of affairs by reference to the interest of those affected by them implies that such states are 'good where they are good for those affected by them'.¹¹²

So, leaving the attempt to save Moore's account to one side, let's consider whether Baldwin's suggestions provides a way of reconstructing the OQA, as Baldwin thinks they do.¹¹³ The first thing to note is that even though this is a naturalistic approach to the concept 'good', it is not itself a reductionist one for:

... the concept of a human interest is an *ethical concept*, linked to judgments concerning the aspects of a life which make that life worth living.¹¹⁴

Thus, for the reductionist to succeed, he would have to demonstrate that he can reduce judgements concerning what 'makes life worth living'. Baldwin suggests that Sidgwick provides

110 Moore (1993: 153).

111 Moore (1993: 153).

112 Baldwin (1990: 74).

113 Baldwin (1990: 74-101).

114 Baldwin (1990: 91). Emphasis mine.

the most plausible attempt at this.¹¹⁵ Sidgwick argues that the aspect of life that makes life worth living concerns 'no more than the intrinsic desirability of aspect of human life.'¹¹⁶ We can modify the details of Sidgwick's account; a good candidate would be Peter Railton's naturalism.¹¹⁷ Railton claims that the aspects of a life that make that life worth living are those which (roughly) our idealised self would desire our non-idealised self to desire.

What, then, is the relationship between this and Baldwin's reconstructed OQA? Well, first notice that Moore would claim that because it is a significant question whether what our idealised self would desire for our non-idealised self to desire is in fact good (or valuable) for us, we can reject Railton's account. However, Moore's argument fails in providing convincing reasons to take his intuitions seriously. Here Baldwin makes a few suggestion as to why we ought to 'go with' Moore's intuitions and why, in fact, Moore's OQA is a genuine argumentative device.

Baldwin's suggestion is that the only reason that we might think Railton's account is plausible is because the second-order desires in his account are typically thought be informed by *ethical* knowledge. So, for Baldwin, second-order desires rely on, for instance, knowledge of the value of dispositions such as kindness, loyalty, fairness, and it is for this reason that we accept that the content of second-order desires provides an account of what makes life worth living. Specifically, in terms of Railton's account it is because 'what our idealised selves would want for us', relies on ethical knowledge. Yet, as Baldwin states:

¹¹⁵ Sidgwick (1963).

¹¹⁶ Baldwin (1990: 91).

¹¹⁷ See for example, Railton (1986a). I discuss Railton's account in detail in chapter V.

Obviously, however, a reductive analysis cannot include this ethical content; but once it is eliminated, most of the point of the distinction between first-order and second-order desires is undermined, and we are returned in effect to the simpler Sidgwickian account [the claim that the aspects of a life which make that life worth living concern no more than the intrinsic desirability of aspect of human life]. To this the Moorean objection that such a position makes it impossible for us to evaluate desires themselves does appear forceful; *it cannot, I think, be denied that closing off this dimension of critical thought appears to us as a loss*. Yet what remains to be understood is what lies behind it, once the naive appeal to our sense of the meaning of ethical language is abandoned.¹¹⁸

Baldwin thinks that if we consider the claim that the aspects of a life which make it worth living, we will recognise that such account is missing something . Given the focus on that which makes life worth living Baldwin claims that the focus on 'thin' ethical terms such as 'good' that Moore proposes is unhelpful and far too restrictive. He proposes that we shift the focus from 'thin' to 'thicker' concepts such as kindness, loyalty, and courage.¹¹⁹ This seems a fair proposal; when we make an evaluation using 'thin' moral concepts, like 'right' or 'good' our answer at least seems to be guided by our 'thick' concepts. It is because some choices exemplify admirable characteristics, like compassion, that we judge them to be a good one.

Accepting this claim, though, would mean that the reductionist now has a far more daunting project. As Baldwin notes, the reductionist cannot merely list the ethical concepts to be

¹¹⁸ Baldwin (1990: 91). Emphasis added.

¹¹⁹ This distinction was introduced by Williams (1985: 128-9). This is discussed again in the next chapter when I consider blame.

analysed. Instead, before he can establish that ethical truths are derivable from a theory incorporating no ethical concepts, he has to characterise the ethical.¹²⁰ This is certainly a coherent project: there is descriptive content to 'thick' ethical concepts, and presumably if the reductionist can isolate this content using non-ethical terms he will have taken the first step in his project. However, influenced by John McDowell's work, Baldwin argues that for the reductionist there seems to be no way of characterising the extension of 'thick' concepts.¹²¹ 'Thick' concepts do not just attach an evaluative label to an otherwise ethically neutral disposition; they carve up the world from a point of view that is 'embedded within the evaluative practices of culture.'¹²²

These points suggest to Baldwin one reason why the OQA is as a challenge to the reductive naturalist. First, 'thin' ethical concepts get their content from 'thick' ones and second, 'thick' ethical concepts are irreducible. Therefore, any proposed reduction of 'thin' ethical terms such as 'good' will seem incomplete, and questions concerning them will seem open.

Notice that even though this helps Moore, again he would be very unhappy with at least two of Baldwin's claims. First, Baldwin's suggestion turns on the ability to define 'good' in terms of 'thick' ethical terms, and of course, Moore thinks that 'good' is indefinable.¹²³ Second, Moore thought that good must be a non-natural property; however, Baldwin's position uses 'thick' ethical concepts and:

120 Baldwin (1990: 92). Also, see my discussion of blame in the next two chapters.

121 Baldwin (1990: 92) and McDowell (1979). We could, of course, at least start to describe virtues in descriptive terms. Kindness might be, for example, described as a disposition to pay extra attention to the welfare of others.

122 Baldwin (1990: 92). This does not mean that to grasp the concepts one must have the values that they describe. All that is required is a sympathetic imagination so that we can see the point of values that we do not endorse. But importantly: '... the exercise of such imagination is not the identification of necessary and sufficient non-ethical conditions for the application of these concepts. It is the assumption of an ethical point of view that we do not altogether share.' Baldwin (1990: 92).

123 See Moore (1993) & Chapter I.

...the use of thick ethical concepts has existential implications: there are no general truths about kindness independent of the existence of people.¹²⁴

Hence, although Baldwin's anti-reductionist OQA concedes too much to the ethical naturalist to be acceptable to Moore, we can sum up Baldwin's position as follows.¹²⁵ Judgements about good are those that determine what people *ought to do*. Hence, given that we can think of all evaluations as assuming an interested point of view, and it is this standpoint of appraisal that provides the criterion by which the things to be evaluated are evaluated,¹²⁶ the standpoint that provides the criterion by which we evaluate something as good amounts to human agents wondering *how* to live their lives. As stated, it is the criterion of people asking themselves what it is that makes life worth living. Therefore, the reductionist has to reduce 'good' to a 'value-free' understanding of the ends of human life. However, Baldwin suggests that it is unlikely that anyone possesses such understanding. After all, we gain an understanding of the ends of human life through self-understanding, which is achieved through social identification and individual commitments which themselves incorporate value judgement. Consequently, he writes:

¹²⁴ Baldwin (1990: 92).

¹²⁵ This summary is developed from a later work: Baldwin (1993).

¹²⁶ Baldwin (1993). I think Baldwin, by talking in this way of an 'interested point of view', has in mind something similar to McDowell's (1979) notion of 'concern' and Wiggins's (1993b) notion of 'interest'. See my discussion of Wiggins in chapter V below.

It is these value judgements which, as things stand, make possible the kind of reflective detachment concerning alleged analyses to which Moore's phenomenological argument [his phenomenological version of the OQA] calls attention.¹²⁷

For Baldwin's suggestions to be convincing a lot more work needs to be done on the notion of 'interest'. For instance, why should we accept Baldwin's claim that 'thick' concepts do not just attach an evaluative label to an otherwise ethically neutral disposition, but carve up the world from a point of view that is embedded within the evaluative practices of culture? No doubt, Baldwin would have a sophisticated response to this and could elucidate the notion of 'interest'. However, it is not my aim to discuss this in detail here. The main worry with Baldwin's suggestion is that it does not demonstrate that a reductionist's project will necessary fail. His claim is, at best, a proposal as to why we ought to expect the open question phenomenon. In order to develop Baldwin's reconstruction of the OQA, one would have to show that *conceptually* there could be no value-free understanding to the ends of human life.

§3. Darwall, Gibbard and Railton's OQA.

Darwall, Gibbard and Railton's (DGR's) suggestion differs from Baldwin's; theirs relies on the possibility that the link between moral judgement and motivation/reason is one of conceptual necessity.¹²⁸ They argue that if this were the correct account, then the intuitions people have

¹²⁷ Baldwin (1993a: xx).

¹²⁸ It is possible to read Baldwin's account as implicitly internalist, and hence similar to DGR's suggestion.

demonstrate that analytic naturalism cannot secure their proposed reductions. DGR suggest this could count as an account of the fact that people find questions open.

In a recent review of modern ethics, DGR suggest:

Moore had discovered not a proof of a fallacy, but rather an *argumentative device* that implicitly but effectively brings to the fore certain characteristic features of 'good' – and of other normative vocabulary – that seem to stand in the way of our accepting any known naturalistic or metaphysical definition as unquestionably right, as definitions, at least when fully understood, seemingly should be.¹²⁹

DGR, like Ball and Baldwin, argue that the most important feature of Moore's OQA is that it highlights the phenomenon that people we judge to be psychologically and linguistically competent do find questions concerning proposed reductions of ethical terms to natural terms significant and non trivial. The force of the OQA arises if we:

... observe that the open question argument *is* compelling for otherwise competent, reflective speakers of English, who appear to have no difficulty imagining what it would be like to dispute whether P [some natural property] is good.¹³⁰

¹²⁹ DGR (1992: 3-4). Emphasis mine.

¹³⁰ DGR (1992: 4).

Agents we believe to be psychologically and linguistically normal do in fact find questions concerning proposed conceptual analysis of value terms 'open'. DGR suggest that it is the philosopher's job to explain this; their own 'tentative suggestion' starts from the realisation that questions of the form 'Is P really good?' can be read as 'Is it clear that, other things being equal, one really ought to, or must, devote oneself to bringing about P?'¹³¹ DGR continue:

Our confidence that the openness of the open question does not depend upon any error or oversight may stem from our seeming ability to imagine, for any naturalistic property *R*, clearheaded beings who would fail to find appropriate reason or motive to action in the mere fact that *R* obtains... Given this imaginative possibility, it has not been *logically* secured that *P* is action guiding (even if, as a matter of fact, we all do find *R* psychologically compelling). And this absence of a... *conceptual* link to action shows us exactly where there is room to ask, intelligibly, whether *R* really is good.¹³²

In fact, this quotation contains two accounts of the OQA. One concerns motivation and the other concerns reason. There is, I argue, a close connection between them. However, in order to present DGR's methodology clearly, I will focus on motivation. In the next two chapters, I will return to each argument and distinguish and develop them.

To understand DGR's suggestion, we have to assume a conceptual link between judgement and motivation. They suggest that, as a matter of conceptual necessity, if agents judge it is right to do

¹³¹ Darwall, Gibbard and Railton (1992). Pg. 4.

¹³² Darwall, Gibbard and Railton (1992: 4). Emphasis added.

ϕ , then they will be motivated to do ϕ , unless they are suffering from some motivational disorder.

Call this position internalism.

Thus, for a successful reduction of moral terms to natural terms, the naturalist has to demonstrate that his account respects internalism. However, DGR point out that the naturalist must account for the fact that we think we can imagine people who, by our commonly acknowledged standards, are psychologically and linguistic competent, but who fail to be motivated by their judgement that a natural property obtains. This, they suggest, means that the naturalist cannot demonstrate that he respects internalism. Moreover, this accounts for the phenomenon that people do find questions concerning proposed reductions open. I suggest that this is a strong challenge to the naturalist in as far as internalism is plausible.

The naturalist who is also an externalist would remain unchallenged by DGR's suggestion. Such a naturalist would argue moral terms are reducible to natural terms, that there is no conceptual connection between moral judgement and motivation, and thus there is no conceptual connection between naturalistic judgements and motivation. Adopting this position means that DGR's observation that we can imagine clearheaded beings who judge that natural properties obtain yet fail to be motivated becomes supporting evidence, rather than a reason to question the plausibility of reductive naturalism.

If internalism is correct then DGR's OQA presents a strong challenge to reductive naturalism, as the naturalist would be committed to the claim that, as a matter of conceptual necessity, naturalistic judgements motivate. But he would then be required to account for the convictions of competent and reflective speakers of English that they can imagine competent clearheaded beings failing to be motivated by their naturalistic judgements.

To such a challenge the naturalist could presumably respond by claiming that rather than taking these convictions as evidence against his position, we are best to account for the convictions in terms of the *complexity* of the link between moral and natural concepts. So, the naturalist would agree with DGR's suggestion, but propose that the only reason that we can imagine competent clearheaded beings failing to be motivated is because we do not fully understand what naturalistic reductions involve because they are so *complex* that if we did, we would realise that there are no such possibilities. Alternatively, the naturalist could show that the convictions arise through *systematic confusion*. For instance, perhaps they are due to the residue of religious beliefs distorting people's understanding of moral concepts. Finally, the naturalist could insist there is nothing in DGR's suggestion to show that all *possible* reductions will fail. I discuss these points in detail in chapters three and four.

§4. Conclusion

Stephen Ball argues that the traditional criticisms of the OQA - that it is circular and invalid - fail. There is no circularity because the criterion for openness is not analyticity, but rather what non-philosophers would say about proposed naturalistic reductions. There is no invalidity because we can assume that after rational reflection native speaker's intuitions are typically not distorted.

For the same reason as Ball, Baldwin and DGR reject the circularity objection. However, unlike Ball, they suggest that one could not merely infer from people's linguistic intuitions that any semantic reduction of moral terms will fail. They argue that we must give an independent account as to *why* people think that the openness of open questions does not rely on any error or

oversight. Baldwin gives two suggestions, both of which rely on the fact that morality requires a certain commitment – what he calls people's interests – in virtue of which people make moral judgments. He claims that it is the analytic naturalist's inability to give a semantic reduction of *this* commitment that explains the openness of open questions.

DGR suggest the reason that people think that the openness of open questions does not rely on any error or oversight depends on the action-guiding nature of moral judgements. Specifically, they claim that people's intuitions undermine the possibility of there being a conceptual link between naturalistic judgements and motivation/reason. Hence, if there were a conceptual link between moral judgments and motivation/reason, this could account for people's open question intuitions. In the next two chapters, I discuss DGR's suggestions in light of Baldwin's claims about peoples' interests.

(III)

MORAL MOTIVATION AND THE OQA

Recall that at the end of the last chapter I suggested that if we could give an account of peoples' linguistic intuitions concerning open questions, then the OQA would be a genuine argumentative device. The explanation I discuss in this chapter concerns the link between moral judgement and motivation. To adequately assess this account, a broader discussion of moral motivation – motivation by moral judgements – is required. If this enquiry concludes that as a matter of conceptual necessity moral judgments are connected to motivation, then Darwall, Gibbard and Railton's (DGR) OQA presents a strong challenge to analytic naturalism.¹³³ On the other hand, if the connection between moral judgement and motivation is wholly contingent, then DGR's OQA fails. I will conclude that it does fail, but argue that the nature of the link between moral judgment and motivation challenges the analytic naturalist.

§1. Moral Judgements and Motivation

In daily life, we encounter people who appear always to act in accordance with their moral judgements. For them, 'doing their duty' is of utmost importance, even if this means that relationships with family and friends suffer. In the other extreme, we meet people for whom there appears to be no relation between their apparent moral judgements and their motivation. For instance, they may agree with us that they have a moral obligation to give to Oxfam, but

¹³³ The direction of the possible necessary connection is from judgement to motivation.

wonder why this should concern them. Such characters' moral judgements invariably fail to motivate them.

However, the majority of people fall somewhere in between these cases. Their moral judgements, for the most part, motivate them; however, if they are disillusioned or depressed their moral judgements fail to motivate.

It seems then that there is *prima facie* reason to think that the connection between moral judgement and motivation is wholly contingent.¹³⁴ In turn, this suggests that moral judgements (in their relation to motivation) are on a par with empirical judgements of fact.¹³⁵

However, traditionally philosophers have suggested that drawing such an analogy with empirical judgements of fact would miss something distinctive of moral judgements as it seems to make the connection between moral judgement and motivation too arbitrary.^{136 137} We can highlight this worry by an example. A friend fails to be motivated to turn off the TV by his judgement that it is showing nothing of interest. We think nothing more of it, except perhaps that he is lazy, whilst we would be puzzled and worried if our friend judged it morally wrong to steal from supermarkets, but carried on taking tins of baked beans from Tesco. As such, it appears both that the link between moral judgement and motivation is not the same as for non-moral judgements. But, as we have already noted, the connection between moral judgements and

¹³⁴ 'Wholly contingent' is meant to capture the thought that an agent may have *never* been motivated by his apparent moral judgement. That is, it rules out a link of logical dependency between moral judgement and motivation. This will become relevant below.

¹³⁵ Of course, non-moral (but still evaluative) judgements motivate; for example, the judgement that this food is disgusting motivates me not to recommend it to someone.

¹³⁶ Both internalists and externalists make this point. See for example, Falk (1986), Smith (1994), McNaughton (1988) and Brink (1997).

¹³⁷ In fact, this is one of my worries about the externalist position: that it appears to make the link between moral judgement and motivation *ad hoc*. See below.

motivation appears wholly contingent. How we deal with such thoughts designates us as either 'internalists' or 'externalists'.

Internalists claim that our intuitions about the link between judgement and motivation being wholly contingent are simply mistaken. There is a necessary connection between moral judgement and motivation. For them, this contrasts with non-evaluative judgements of empirical fact.

Alternatively, externalists claim that we should just 'look at the world' and recognise that the link between moral judgement and motivation is wholly contingent. They say that peoples' thoughts about the 'special' nature of the link between moral judgement and motivation reveal only a purely contingent psychological fact. For the externalist, our puzzlement concerning our friend's lack of motivation to stop taking cans of baked beans arises because it just so happens that typically, people desire to do what is right. Nevertheless, the externalist insists that things could have been completely different. People could normally be unmoved by their moral judgements.¹³⁸

Things become complicated quickly. This is because philosophers use 'internalism' to refer to a host of different ideas. Within meta-ethics, the only thing common to all accounts is that they attempt to clarify the link between moral judgement and motivation.¹³⁹ The internalism I discuss claims that there is a *necessary* connection between moral judgements and motivation,

¹³⁸ Maybe to illustrate the difference the externalists could reconstruct a (perhaps unbelievable) scenario in which we have all come to desire to stop watching uninteresting TV, but where we have all lost the desire to do what is right. In which case, our puzzlement expressed in the example given above, would be completely reversed. We would be worried about our friend's approach to TV watching, not to him stealing his lunch.

¹³⁹ See for instance, Darwall (1983), Jackson (1988), Smith (1994), for different versions of internalism.

and importantly this is *conceptual necessity*.¹⁴⁰ If someone apparently makes a moral judgement yet fails to be motivated, then the internalist claims that he or she cannot be credited with moral judgements. The best we can say is that he or she only *appears* to make moral judgements.¹⁴¹ Internalism provides ascription conditions for moral judgements in terms of moral motivation.

However, philosophers argue – I suggest correctly – that this makes the connection between judgement and motivation unacceptably strong.¹⁴² For example, if I made a moral judgement when in good health, and the reason for me making such a judgement remain unchanged, (as far as I know/believe) then we might say the judgement is genuinely my judgement even when I become depressed and am not motivated.¹⁴³ It is not *a priori* that if a moral judgement fails to motivate, then it fails to be genuine. For this reason I understand ‘internalism’ as a weaker claim. Moral judgements are of conceptual necessity connected to motivation *except* in individuals suffering from motivational disorders that affect them more generally.¹⁴⁴

Internalism then allows the ascription of moral judgement to agents, even if they are not motivated. Yet, internalists claim that it is *conceptually* impossible for an agent who has no

¹⁴⁰ McDowell (1978) holds such a position.

¹⁴¹ Of course this is just a crude way of putting Hare's (1952) inverted commas account of moral terms. See Smith (1994) for a discussion of this. The idea is that the agent who is not motivated by his moral judgment is not expressing a genuine moral judgement, but rather an opinion about what kind of moral judgment his fellows would make in his circumstance.

¹⁴² There are notable exceptions; for example, McDowell (1978), and Platts (1981), who think that this is acceptable.

¹⁴³ That is, it matters when the judgement is formed. One might question whether a severely depressed person can make a judgement; might it not be devoid of conviction, hence not be a judgement? Alternatively, might the agent be unable to consider the reasons for it, so again it might not be a genuine judgement?

¹⁴⁴ I presume that the ‘without motivational disorder’ clause does not make internalism trivially true. Also this allows for akrasia; the moral judgement does motivate but does not win.

motivational disorders to make a moral judgement and fail to be motivated. It is *a priori* that we cannot ascribe moral judgement in such cases as these.

The debate between internalism and externalism is essentially a dispute over the possible existence of the amoralist - the person who is completely unmoved by moral considerations. The internalist insists that amoralists are *conceptually* impossible while the externalist suggests that everyday life confirms their existence.¹⁴⁵ However, philosophers have recognised that just focusing on the amoralist makes the debate intractable. As such, they suggest that the dispute between the internalist and the externalist should start from a claim that both accept. From this one side could demonstrate either that their account fits better with such a claim, or that the other side is incompatible with it. It is for this reason that Michael Smith introduces the 'striking fact' that a change in motivation follows a change in moral judgement. He argues that both internalism and externalism would accept this fact, yet it is only compatible with internalism. However, I want to consider another persuasive claim put forward by Sirgún Svavarsdóttir who arrives at a different conclusion.¹⁴⁶ She states that both internalists and externalists would *accept* a certain methodological principle, but only the externalist can *respect* it. To demonstrate this she starts with an imagined scenario:

Virginia has put her social position at risk to help a politically persecuted stranger because she thinks it is the right thing to do. Later she meets Patrick, who could, without any apparent risk to himself, similarly help a politically persecuted stranger, but who has

145 For fascinating account of an alleged 'real life' amoralist see Watson (1987).

146 Svavarsdóttir (1999).

made no attempt to do so. Our morally committed heroine confronts Patrick, appealing first to his compassion for the victims. Patrick rather wearily tells her that he has no inclination to concern himself with the plight of strangers. Virginia then appeals to explicit moral considerations: in this case, helping the strangers is his moral obligation and a matter of fighting enormous injustice. *Patrick readily declares that he agrees with her moral assessment, but nevertheless cannot be bothered to help.* Virginia presses him further, arguing that the effort required is minimal and, given his position, will cost him close to nothing. Patrick responds that the cost is not really the issue, he just does not care to concern himself with such matters. Later he shows absolutely no sign of regret for either his remarks or his failure to help.¹⁴⁷

Importantly this description is neutral on whether Patrick is an amoralist; the externalist would claim he is, and the internalist would deny this on the ground that he is not really making a moral judgement. However, this means that as it stands the example will resolve nothing; we arrive at conflicting intuitions. Until we discuss more detail the externalist cannot claim that Patrick makes a moral judgement, and the internalist cannot claim that he does not. Svavarsdóttir provides some neutral additional information that she suggests favours externalism.

Additional information about Patrick: Besides being known for courage and conservative estimates of risk to himself, Patrick is independently minded and earnest to a fault –

¹⁴⁷ Svavarsdóttir (1999: 176-7). Emphasis added.

indeed, honest to the point of tactlessness and even cruelty. And in any case, he has nothing to gain from misleading Virginia in the given circumstances. Moreover, Patrick makes claims couched in moral terms infrequently and impassionately, and never gives them as reasons for his actions. He has frequently been observed taking actions that seem uncontroversially wrong... without displaying any signs of hesitation or regret. In contrast, he has often displayed obvious signs of regret and shame when his plans have misfired, he has overestimated risk to himself, or he has publicly embarrassed himself in matters he finds important. He has also passed up numerous opportunities to perform obvious and uncostly moral deeds. However, when prodded, he will engage in prolonged and intelligent conversations about moral matters and seemingly take an independent stand on the moral statues of a controversial public policy or an action. Nonetheless, he usually ends such conversations by volunteering the opinion that he has long ago rid himself of any aspiration to live by moral standards.¹⁴⁸

Svavarsdóttir thinks that with this additional information the best way of describing Patrick is as follows: he makes a moral judgment but fails to be motivated because he lacks the appropriate conative attitude. Hence, she thinks that this additional knowledge about Patrick favours externalism. However, the internalist will remain unconvinced; after all, we can interpret the example (including the additional information) as Patrick only *appearing* to make moral judgments. The question then turns on whether externalists can stop such a redescription. Svavarsdóttir argues that they can. Presented with an example of someone such as Patrick - where someone makes a moral judgment but remains

148 Svavarsdottir (1999: 177-8).

unmotivated - an internalist would reject the example as incoherent. But, in so doing, he is ruling out a plausible explanation of this phenomenon, namely, the externalist's. For instance, the externalist may argue that Patrick fails to be motivated by his judgment because he lacks the appropriate desire. However, it is a reasonable methodological principle to require those who would restrict the available hypotheses for explaining a given phenomenon to carry the burden of argument for this restriction. But internalists have not, so far, typically done this, or where they have tried, they fail. I think this is a fair criticism of internalism. Moreover, it fits with my intuition that it is not *a priori* that a psychologically competent agent whose moral judgement fails to motivate him did not, in fact, make a moral judgement. Internalism then is in trouble.¹⁴⁹ Moreover, given that DGR's argument rests on internalism, it is also in trouble. Yet, I suggest that the link between moral judgement and motivation may still cause problems for the analytic naturalist.

Given I reject internalism, does this mean that I accept externalism? If the externalists are mainly interested in defeating the internalist thesis that there is a conceptual connection between moral judgment and motivation, then yes. However, if their challenge to internalism is that moral judgement cannot motivate in and of themselves, then no.¹⁵⁰ I think it best that I reject the label 'externalist', as typically in the literature externalism includes both these features.¹⁵¹ However, why don't I agree with both?

149 There are other worries I have concerning internalism. One is this: Michael Smith claims that a change in judgement brings with it a change in motivation. He states that this is a 'striking fact' that the internalist would accept, and only he could account for. However, it is controversial whether a conceptual truth could account for such a counter-factual claim (see Jackson (1987)). It then turns out - ironically - that it is the internalist, rather than the externalist, who has problems accounting for the 'striking fact'.

150 They think this would challenge internalism, because they typically think that motivation only takes place if there is belief and desire, and judgement lacks any desire. See below.

151 See for example, Brink (1997) and Railton (1993a).

One of the main reasons why externalists think that moral judgement cannot possibly motivate in and of itself is because they typically (though not always) start from the presumption that belief-desire psychology is correct.¹⁵² This is the view that an intentional action must be describable in terms of an appropriately related belief and desire had by an agent. For example, if I went to retrieve the post, and my desire is to read the post, then on this account my belief must be that going to get the post will satisfy my desire to read the post.¹⁵³ The account comes from David Hume, and recently philosophers have defended it.¹⁵⁴ The externalist accepting this theory would claim that I am right in rejecting the conceptual connection between moral judgement and motivation, but wrong in thinking that moral judgment can motivate alone. For judgements to motivate there needs to be a further desire. Externalists differ on how to account for this desire; it is often thought of as a desire taken towards objects under a *moral* mode of presentation, specifically, *the desire to be moral*.¹⁵⁵

However, I reject belief-desire psychology. I do not think that all intentional action must be explicable in terms of appropriately related beliefs and desires.¹⁵⁶ However, in rejecting this position I am not claiming that an account of motivation in terms of the belief-desire theory is

152 Notice, as I discuss below, even though not all externalists explicitly start from this presumption, all I have encountered arrive at thinking which presupposes it. See for example, Svavarsdóttir (1999).

153 Notice that the internalist who is committed to the belief-desire theory has a number of immediate problems, one of which is the fact that it would commit him to ascribing desires in cases where we would not want to. For instance, say I found my daughter committing a crime. I judge it is right to give her over to the police, although I know this will mean a heavy prison sentence. If I am an internalist, and am committed to the belief-desire account, then I *necessarily* have the desire to give my daughter over to the police. Surely, though it seems odd to think that because I make the moral judgement, I have, of necessity, a non-derivative desire which accords with the belief expressed in my judgement.

154 Hume (1975) & (1978). For example, see Smith (1994).

155 See Smith's (1994) discussion of externalism. He claims that this way of characterising externalism (in terms of the desire to be moral) is the most plausible version, as it is only with such a general (*de dicto*) desire that the externalist can guarantee that moral motivation tracks changes in moral judgement. NB. Svavarsdóttir thinks that we must posit a desire to be moral to account for moral motivation. However, as I noted above, this is not her starting point. She claims that this is best way of accounting for the fact that moral judgment sometimes fails to motivate.

156 I do not discuss my reasons. However, my first attempt at justifying this claim would be that the account is not falsifiable, which counts against it being a good hypothesis.

always false. The point is that, to progress in the philosophy of moral psychology, we must not allow the belief-desire account to become the presumption in virtue of which we redescribe all motivation.

This means, for instance, that there is no obvious reason justifying the inference from the fact that Patrick can make genuine moral judgements yet remain unmoved, to the claim that he must be lacking the appropriate *desire*. In general, one cannot simply assert that because moral motivation is affected by factors other than moral judgement, moral motivation must be grounded in certain conative attitudes: for instance, the desire to be moral.

I reject attempts to explain failures in motivation (absent pathological disorders) by saying 'really there is no moral judgement'. In addition, I reject as sheer dogma, not merely the belief-desire theory, but the view that there must be an explanation (at the intentional level) of motivational failure. I am not required to explain failures in motivation at the intentional level at all, hence *a fortiori*, in terms of a desire to be moral. Moreover, I think that it is not elliptical to claim that we are disposed to do the morally obligatory thing when we take it to be the morally obligatory thing. Specifically, we do not need to postulate further intentional states that ground our dispositions. Conceiving of something in moral terms is enough to be motivated. One can adopt an empirical thesis that captures the attractive features of internalism without being an internalist.

One reason this option is perhaps unobvious is due to how philosophers describe amoralists. They write things like: the amoralist could not 'care less', or that he is not 'bothered' with moral matters, or that he is 'left cold' by his moral judgments. All these invite the thought that, if failure in motivation is due to the lack of some conative attitude, then agents must be

motivated by this conative attitude in conjunction with the appropriate belief. However, this obviously does not follow (at least as long as one rejects the belief-desire psychology). Maybe it is true that one can account for the fact that moral judgement can fail to motivate in terms of the lack of certain desires, but this cannot demonstrate that the reason moral judgements *do* motivate is due to the presence of certain desires. This will become clearer in the following discussion. We can reject internalism without having to adopt such a move once it is recognized that there are many indicators of whether a person possesses a moral concept and makes a moral judgment. I demonstrate this using the indicator of blaming others.

Consider again the example of Patrick. Assume he *does* see what is right but fails to be motivated. This, the internalist would insist, is not merely puzzling but is conceptually impossible. True, it is puzzling: surely, though, it is not conceptually impossible. Can we know *a priori* that any agent who fails to be motivated by their moral judgement and who is suffering from no motivational disorders, did not make a genuine moral judgment? It is doubtful. As I stated there are surely other indicators of whether the person makes a genuine moral judgement.

Supposing that he makes a genuine moral judgement, imagine we ask Patrick whether he thinks there is a reason to blame him for not helping the politically persecuted strangers. If he says no, then either we will think that he is lying, or we will think that he really fails to grasp what it is to have an obligation. The reason for the latter claim is the obvious truth that if I fail to do what I have moral obligation to do then I am blameworthy.¹⁵⁷

¹⁵⁷ I discuss this truth in more detail in the next chapter.

It follows then that if Patrick genuinely believes he has a moral obligation to help the politically persecuted stranger, then he will think that he is blameworthy for not helping him.¹⁵⁸

Furthermore, we can claim this because it is a platitude that he cannot be blameworthy if he did what he had most reason to do. As such, he must think that the contraposition of this platitude is true, namely, if he *is* blameworthy then he has failed to do what he has most reason to do. Consequently, from Patrick accepting that he has a duty to help, he accepts that he has most reason to help the strangers. As Skorupski writes, Patrick ought to reason as follows:

As Virginia says, *I have a duty* to help these politically persecuted strangers. So it's quite reasonable to blame me for not helping them. But it wouldn't be reasonable to blame me for not helping if not helping was what I had most reason to do. So what I have most reason to do is help.¹⁵⁹

However, even granted that there is such a link between moral judgement and practical reason, this still does not address motivation.¹⁶⁰ I think that *the judgement that something is what one has the most reason to do has the potential to motivate in and of itself*. There is no need to posit a further desire to do what one has reason to do for the judgement to motivate. Again, the temptation to do so may arise from the belief that in cases where such a judgement fails to motivate, we need to give an account at the intentional level. However, there is no reason to think this. Specifically, there is no reason to think either that there is always an absent single

¹⁵⁸ Obviously he does not have to *feel* guilt to make this judgement. Perhaps, though, it is necessary that he *has* experienced guilt. See the next chapter for more on this issue.

¹⁵⁹ Skorupski (1999b: 3) Emphasis added.

¹⁶⁰ I discuss this connection further in the next chapter.

unifying desire, or that we need to explain failed motivation at the intentional level. Consider an example from John Skorupski illustrating this.¹⁶¹

I am driving down the road and notice that the fuel gauge is low. I think to myself 'I should stop at the next petrol station', and immediately I see a turn-off for one.¹⁶² In scenario A, I pull in to the garage, and in scenario B, I fail to slow down and drive straight past. In the latter case, there seems no reason to demand an account of why motivation failed that applies to all cases of failed motivation, nor does there seem to be any reason for demanding an explanation at the intentional level.¹⁶³ For instance, in scenario B I may have low blood sugars due to spending hours driving.

I have demonstrated that moral judgements can motivate in and of themselves because of their link to judgements about what there is most reason to do. Hence, moral judgements can motivate in and of themselves, but they need not do so. In the example of Patrick, even if he admits he had most reason to help the strangers, he may not be motivated to do so.

We could write more about Patrick. We could claim that even though someone did what he or she had most reason to do, there is still reason to blame him or her for doing it. Agreeing that he has a duty, and that it is reasonable to blame him for not doing it, he could deny that this means he should help the strangers. Here we may start to wonder if Patrick really understands what it is for something to be blameworthy.

In this section, I have put forward three theses. First, I have argued that the amoralist is not a conceptual impossibility. It is not *a priori* that we can only ascribe moral judgments to

¹⁶¹ See Skorupski (1999b: 2).

¹⁶² By the claim that I 'should' do x, I mean that I have the most reason overall for doing x.

¹⁶³ For example, one may think that the reason for my failing to be motivated in B must be due to the lack the desire to do what I take myself to have the most reason to do.

agents who are motivated by their apparent judgments. Second, it is empirically true that moral judgments can and do typically motivate in and of themselves.¹⁶⁴ Third, this is due to a link between moral judgements and what one has the most reason to do. Furthermore, I have rejected belief-desire psychology as a starting point in the discussion of motivation. Hence, I think that the amoralist does not demonstrate that moral motivation must be grounded in a conative attitude. It is true that factors other than the judgements themselves may affect the motivational impact of the judgement. However, this does not mean the motivation is best accounted for in terms of belief-desire accounts.

I am neither an internalist nor an externalist as such types are usually characterised. I reject externalism as far as it accepts belief-desire psychology; I reject internalism for its claims about conceptual necessity. The pressing question is whether this account renders the OQA a genuine argumentative device. Does the fact that judgements about one's moral obligations *in and of themselves* typically, though not necessarily, motivate, render the OQA a genuine argumentative device?

§2. Motivation and DGR's OQA.

Recall, that DGR's suggestion ran as follows.¹⁶⁵ First, we ought not to claim utter conviction about open questions, but note that the open question argument is compelling for otherwise competent, reflective speakers of English.

¹⁶⁴ I suspect that this is in fact all that the internalist wants, but attempts to capture it by the impossibly strong account of conceptual necessity.

¹⁶⁵ DGR (1992).

Second, one should articulate a philosophical account of why this might be so. DGR's suggestion is that our confidence that the openness of open questions does not depend upon any error or oversight stems from our seeming ability to imagine linguistically and psychologically competent beings judging that natural properties obtain but not being motivated. DGR claim that given this, it has not been secured that there is a conceptual link between naturalistic judgements and motivation.

Naturalists – given what I said in the last section – will be unimpressed by this. They could agree that there is no conceptual link between judgements about natural properties obtaining and motivation, yet point out that there is also no conceptual link between moral judgements and motivation. Hence, this is not an objection to the reduction of moral terms to natural terms. DGR's suggestion fails.

Even given this failure, there is perhaps a challenge to analytic naturalism arising from the last section. That is, can analytic naturalists demonstrate that naturalistic judgements have the capacity to motivate in and of themselves? If they cannot then this could count as an account for why people find open questions 'open'.¹⁶⁶

Consider a paradigm example of a natural judgment which naturalists may give: the judgement that something causes pain. If they are asked why such a judgement motivated, their likely response would be that it was *terrible*, *horrible*, or *abhorrent*. But then it is only in virtue of an *evaluative* characterisation of the pain that such a fact has the potential to motivate in and of itself. Other likely candidates for naturalist judgements that have the capacity to motivate are

¹⁶⁶ One could also ask: is there any natural property N, such that I am blameworthy if I judge that an action has N and am not moved to do it (or just fail to do it)? I discuss this in the next chapter.

'x would be pleasant for me' and 'x is in my interest'.¹⁶⁷ Arguably, there need be no distinct desire for having either of these properties. However, both types are very poor candidates for being the reductive basis for 'I morally ought to pursue x'; that is, given that we continue to believe that sometimes we morally ought to assist strangers by actions not at all pleasant to ourselves. I think the burden of proof is on the analytic naturalist to come up with an example of a natural predicate that has more or less the same extension as its supposed normative equivalent and is such that judgment formed with it motivates, much of the time. In addition (as I will discuss in the next chapter), the predicate must be such that most people who make judgements about themselves while employing that predicate feel guilty and expect to be blamed if they don't act accordingly. It seems to me not enough for analytic naturalists that they can imagine a world in which everyone has been conditioned to pursue, without an additional desire for it, the greatest good of the greatest number, and conditioned also to feel guilty, dish out blame etc. accordingly. How can this imagined possibility show us what 'ought' does actually mean, as distinct from showing us that there must be a world in which it does mean 'leads to the greatest good'. To me, it is unclear whether naturalists could ever demonstrate that naturalistic judgements could have the capacity to motivate in and of themselves.

Hence, they must either reject the claim that the disposition to be motivated by naturalistic judgments must be grounded in a desire or they must demonstrate that, contrary to what I have argued, the disposition to be motivated by one's moral judgement must be grounded in desire. Perhaps they could do this via a defence of belief-desire psychology.

¹⁶⁷ Presuming the 'interest' can be thought of as a natural predicate.

§3. Conclusion

I have argued that it is incorrect to think that the ascription conditions for moral judgments are such that an agent could not be considered to have made a specific moral judgement unless he were motivated in a specific way. There are, I suggested, other indicators of whether an agent has made a moral judgement. Specifically, if an agent fails to recognise that if something were his moral duty then he is blameworthy for not doing it, then we would doubt he really understands what it is to have a moral duty.

I suggest that because of the relationship between moral judgment and blameworthiness we can accept that moral judgements have the capacity to motivate in and of themselves.¹⁶⁸ This is because the judgement that *x* is what I should do can, in and of itself, motivate me. And, the disposition to do what I have most reason to do need not be grounded at the intentional level at all, *a fortiori*, in the desire to do what I have most reason to do.

All this means that even though DGR's suggested reading of the OQA fails as it relies on there being a conceptual link between moral judgement and motivation, there is a possible challenge to analytic naturalism. That is, given the truth of this chapter, the challenge to the analytic naturalist is whether they can demonstrate that naturalistic judgements, in particular, the naturalistic judgments to which they say moral judgements reduce, have the capacity to motivate in and of themselves.

In the next chapter, I argue that there is a far stronger challenge to analytic naturalism. This develops from DGR's second suggestion concerning the OQA: namely, that there is a

¹⁶⁸ See the next chapter for a discussion fending off the objection that 'moral' cannot be explained in terms of 'blameworthy' because 'blameworthy' implies 'morally ought to be blamed'.

conceptual connection between moral judgement and practical reason that accounts for why we think the openness of open questions does not rely on any error or oversight.

(IV) THE OQA AND RATIONALISM

In the last chapter I dealt with one of DGR's suggestions concerning the OQA. In this chapter I discuss another suggestion concerning the relationship between moral judgement and practical reason.

§1 considers Smith's rationalism; §2 discusses the following claims: (i) If it is right for x to ϕ then x is blameworthy for not ϕ -ing; (ii) if x is blameworthy for not ϕ -ing, then there is a reason to ϕ ; thus, (iii) if ϕ -ing is right then there is a reason to ϕ . I will argue that with certain qualifications we should think of these as conceptual truths. §3 applies the conclusions of §2 to DGR's OQA. §4 discusses the conclusions of the last two chapters in relation to Frank Jackson's analytic naturalism.¹⁶⁹ §5 draws some conclusions.

§1. Smith's Rationalism.¹⁷⁰

Clearly there is a close link between moral judgement and practical reason. It would be puzzling if I judged it right to give money to charity and, in the same breath, demanded a reason to give money to charity. The central question in this chapter concerns the nature of this link. In particular, I am interested in the status of this claim:

(M) if ϕ -ing is right then there is a reason to ϕ .

¹⁶⁹ I focus on Jackson (1998).

¹⁷⁰ Smith (1994). Chapter 3, specifically §3.9.

DGR's suggestion concerning the OQA depends on one being able to recognise the truth of (M) simply and solely by virtue of grasping the concepts involved; that is, (M) must not be elliptical. It cannot require that we add a further conditional to the antecedent concerning, for example, the agent's desires. Michael Smith has recently suggested an argument that supports this claim, and hence would sustain this version of the OQA.¹⁷¹ However, it is unclear and ultimately fails to illuminate (M).

Smith starts from the observation that we approve and disapprove of what people do when moral matters are at stake. Imagine that I hear that a friend borrowed some money, promising to pay it back, but he does not. In this instance, I am likely to disapprove of his actions. Alternatively, it is more appropriate for me to say that I dislike, rather than disapprove, of him eating chocolate ice cream. We can understand the difference between these cases in terms of me having a legitimate ground for expecting my friend to pay back his debt, but having no such ground to expect anything about his ice cream eating; there is no shared standard in terms of which we can judge ice cream eating. Hence, Smith claims that approval and disapproval are only ever in place when there exists grounds for legitimate expectations about how someone will behave. Smith continues:

But now, in light of our account of the preconditions for approval and disapproval, an obvious question presents itself. Consider, for instance, disapproval of those who act contrary to moral requirements. Such disapproval is ubiquitous. Yet how can this be? For,

¹⁷¹ Smith (1994: section 3.9). Note that Smith does not link (M) to the OQA. Also, notice that he calls the commitment to (M) 'rationalism'. However, given the link to sentiment in (i) and (ii), I reject this label.

as we have seen, disapproval for those who do not do what they are morally required to do presupposes the legitimacy of our expectation that they will act otherwise; it presupposes that, as we see it, their decision is a *bad* one in terms of the commonly acknowledged standards by which their decisions are to be judged. But what provides grounds for the legitimacy of this expectation?¹⁷²

Smith thinks that the only plausible answer demands that we accept (M). He first points out that it is implausible to think that what grounds the legitimacy of this expectation is the fact that rational creatures have entered into an agreement to act morally, for the simple reason that people have never made such an agreement.

Smith's suggestion is that what legitimates expectation is a hypothetical agreement that people would make *if* they were rational. Hence, Smith states that what legitimates our expectation that people will act morally is not an agreement, but the very fact that people are rational agents:

Given that moral approval and disapproval are ubiquitous, the truth of the rationalists' conceptual claim thus seems to be entailed by the fact that the preconditions of moral approval and disapproval are satisfied.¹⁷³

¹⁷² Smith (1994: 89). Emphasis added. Smith here appears to invite a normative reading of 'expectation', that the quotation should be read in terms of what we '*ought* to expect'. However, if this is the case then Smith appears to be begging the question. I discuss this below.

¹⁷³ Smith (1994: 90).

This argument from Smith is unconvincing. If 'expectations' is read non-normatively, then the desired conclusion does not follow, whereas if they're read normatively, the question is begged.¹⁷⁴

Consider an example. My friend is a habitual thief. Suppose I disapprove of his actions. For Smith this is explained in terms of me having a legitimate expectation that my friend will not steal. However, this cannot be the expectation that we will not steal. I know that he is a habitual thief. Thus, for Smith's account to be plausible we must read 'expectation' in a normative way rather than a descriptive one. Hence, on this reading all that we require in order to satisfy the precondition of moral approval and disapproval is legitimate moral 'oughts'. Now, if we presume that the legitimacy of moral 'oughts' entails that our concept of a moral requirement is a concept of a reason for action then Smith's conclusion is vindicated. However, Smith cannot simply help himself to this claim as this is precisely what he is attempting to show. He cannot simply claim that we can satisfy the precondition of moral approval and disapproval because moral 'oughts' entail that our concept of a moral requirement is a concept of a reason for action.

To justify rationalism Smith assumes that our concept of moral requirement is a concept of a reason for action. However, even if Smith's argument can avoid such worries, I do not know why Smith thinks it would illuminate (M).

§2. (M) and Blame.

¹⁷⁴ Smith writes in a number of places that he intends such a normative reading. For example, Smith (1994: 89).

Smith is correct to focus on moral attitudes when discussing the connection between moral judgement and practical reason. However, while he focuses on the general 'approving and disapproving', I – developing the theme of the last chapter – concentrate on *blame*. In this section, I argue that properly understood (M) is a conceptual truth. It is a truth that we can see simply and solely in virtue of grasping the concepts involved.

If I judge that my friend is morally wrong to break his promise then this amounts to blaming him.¹⁷⁵ It is to claim that his action is blameworthy; that is, that there is reason to blame him for breaking his promise.¹⁷⁶ Blame is characterised by a *feeling*; hence, in judging that my friend was morally wrong in breaking the promise, I judge there is reason to *feel* blame towards his action. However, just as I can judge that there is a reason to feel sorry for my action, but not feel sorry, I can judge that there is reason to feel blame without feeling blame.

True, the nature of the feeling is not clear. For instance, it is not obvious that blame has a distinct phenomenology. However, this is unimportant for the rest of the discussion. I assume that whatever the feeling amounts to, blaming is an act or attitude whose notional core is a feeling, in the same way that feeling sorry is the notional core of an apology.¹⁷⁷

Moral judgements can be characterised by evaluative reasons concerning blame. Moral propositions are propositions about what there is reason to *feel*.¹⁷⁸ The morally wrong is that towards which it is reasonable to feel the sentiment of blame. The morally right is that which it

¹⁷⁵ We have here to take into account certain extenuating circumstances, see my discussion on pg. 79 concerning (M*).

¹⁷⁶ There is a wider notion of blame which is used in relation to causes. For example, one may blame a loose nut on the railway track for the accident. However, these types of cases can be distinguished from the ones under consideration because they have no feeling core. In blaming the nut for the accident I am not claiming that it is reasonable to have any *feeling* towards the loose nut.

¹⁷⁷ For an extended discussion of the 'blame-feeling' see Skorupski (1999a: Paper VII).

¹⁷⁸ Note that they are still truth-apt propositions and hence I reject noncognitivism.

would be reasonable to feel the sentiment of blame towards not doing. Other moral propositions could be derived in a similar way.¹⁷⁹

Notice that although we can *characterise* morally wrong in terms of that towards which it is reasonable to feel blame, we cannot *define* 'morally wrong' in terms of blame; in fact I doubt that we can define 'morally wrong'.¹⁸⁰ Even so, this does not mean that the concept that it expresses is simple. The concept is complex as it is constituted by practical reason and that which it is reasonable to blame.

To define 'morally wrong' in terms of the blameworthy would be circular. The concept of moral wrongness enters into the intentional content of blame-feeling. Blame-feeling is blame-feeling *that...* where '...' includes the belief that something was done that *morally* ought not have been.¹⁸¹ My blame feeling towards my friend's promise breaking is a blame feeling *that* my friend has broken his promise but morally ought not to have done. It is for this reason that Skorupski writes:

The blameworthy and the morally wrong are indeed equivalent but the semantically prior term is 'morally wrong'. 'Blameworthy' means something like 'meriting that evaluative response which is appropriate to the morally wrong', and so it is circular to define 'morally wrong' in terms of 'blameworthy'.¹⁸²

179 See Skorupski (1997c: 30).

180 I suggest this is also true of other moral terms. See my discussion of 'ethical nominalism' in chapter IX.

181 See Skorupski (1999a: 31).

182 Skorupski (1999). Pg. 31.

That which it is reasonable to blame partially constitutes the concept of morally wrong. If ϕ -ing is morally wrong, then there is a reason to blame me for ϕ -ing. We can recognise this as true simply and solely by grasping the moral concepts involved. It is a conceptual truth.

However, we need to answer a further question: What is the relationship between ϕ -ing being morally wrong and there being *reason not to ϕ* ? Put bluntly: what has blame got to do with (M)?

I argue that in grasping the concept of blame we recognise that if someone is blameworthy for doing ϕ , then there is reason not to do ϕ : this is a conceptual truth. Imagine how puzzling it would be if I blamed my friend for not keeping his promise but also claimed that there is no reason for my friend to keep his promise. As Skorupski writes:

One simply doesn't count as giving reasons for blaming a person *unless* one is giving reasons why that person should not have done what he did. If I can establish that there was no reason for a person not to α I *have*, unconditionally, established that that person was blameless in doing α . I do not have to establish that any further condition obtains...¹⁸³

Hence, there is a conceptual connection between someone being blameworthy for not doing ϕ and there being a reason to ϕ . Those who would deny its status as a conceptual truth fail to grasp the concept of blame.

¹⁸³ Skorupski (1999). Pg. 43.

Thus, to sum up: what is morally wrong is that towards which it is reasonable to feel the sentiment of blame. This is a conceptual truth. Furthermore, if there is reason to blame someone for ϕ -ing then there is a reason not to do ϕ ; this too is a conceptual truth. One recognises both these claim as true solely and simply by virtue of grasping the concepts involved. By transitivity then, I conclude that it is a conceptual truth that if ϕ -ing is right then there is a reason to ϕ ; (M) is a truth we can recognise simply and solely by virtue of grasping the concepts involved.¹⁸⁴

I have purposely oversimplified a number of issues here. For instance, it is not clear that one can blame an agent for doing what is morally wrong when he/she cannot recognise it as such. If someone cannot recognise murder as wrong, then can we really say that he or she is blameworthy for murdering? Perhaps we cannot. Furthermore, if an agent has an obligation to ϕ yet he has no way of ϕ -ing, then arguably we would not think he is to blame for not ϕ -ing. For instance, consider a friend who fails to meet you as promised because he is stuck on the train. It seems that he is not to blame even though he has failed to discharge his obligation to you.

For this reason I suggest we modify the initial claims.¹⁸⁵ If ϕ -ing is right *and* x is able to tell it is *and* is capable of ϕ -ing, then x is blameworthy for not ϕ -ing. Furthermore, it is a conceptual truth that if x is blameworthy for not ϕ -ing then there is a reason to ϕ . Consequently, I claim that a modified version of (M), (M*) is a conceptual truth.

(M*) if ϕ -ing is right, x can tell it is and x is capable of ϕ -ing, then there is a reason for x to ϕ .

¹⁸⁴ Some philosophers have denied these claims, for instance, Svavarsdóttir (1999). For a brief discussion of this see the next section.

¹⁸⁵ Whilst keeping their status as conceptual truths.

In the next section, I suggest that the OQA presents the analytic naturalist with a challenge: given peoples' linguistic intuitions, how can a semantic reduction respect (M*)? To do this, the analytic naturalist must demonstrate either that the OQA is compatible with (M*), or argue that we should reject (M*).

§3. DGR's OQA and Rationalism.

A challenge using (M*) will be stronger than the one considered in the last chapter. This is because while there is no conceptual necessity between moral judgement and motivation, there is between something being right and there being reason to do that thing, and we are inclined to think that a conceptual truth should be reflected in peoples' linguistic intuitions.

Consider then such a version of the OQA. First, we ought not to claim that open question demonstrate anything beyond doubt, but start from the fact that people do find questions concerning proposed semantic reductions of moral terms open. Second, we must give a philosophical account of why. Here is one such an account.

Assume that (M*) is correct. It is a fact that competent and reflective speakers of English are convinced that they can imagine psychologically and linguistically competent people who judge that they have no reason to ϕ by virtue of the fact that ϕ -ing would produce natural property N .

If it is not a conceptual truth that if ϕ -ing would produce natural property N then there is a reason ϕ , then competent and reflective speakers of English should have such a conviction about

psychologically and linguistically competent people. But then, unless there is some other account as to why competent and reflective speakers of English have the convictions they do, we can conclude that there is no conceptual connection between judging that ϕ -ing would produce N and the judgement that they have any practical reason to ϕ . Consequently, in the absence of another account as to why competent and reflective speakers of English have such convictions concerning psychologically and linguistically competent people, we are not entitled to conclude that judging that ϕ -ing would produce N is the moral judgement that ϕ -ing is morally right. Therefore, the analytic naturalist has not secured the truth of his account.

The analytic naturalist could give a number of possible responses to this. First, he could reject (M*). If it were false, then there is no challenge. To do this the naturalist could deny that either or both of these claims are conceptual truths. (i) If ϕ -ing is right and x is able to tell it is and is capable of ϕ -ing, then x is blameworthy for not ϕ -ing; (ii) if x is blameworthy for not ϕ -ing then there is a reason to ϕ . In fact, I think that there is a difficulty in defending (i) and (ii). The problem is that they are truths that people are spontaneously inclined to assent to, and trying to justify such spontaneous convergence makes their truth seem less plausible. This then presents an immediate problem for those defending (i) and (ii). I suspect this is one of the reasons that, even if successful, Smith's account will fail to illuminate (M). However, this will not satisfy the analytic naturalist so I consider in more detail what denying these claims would amount to.

The analytic naturalist may reject (i), claiming that it is not incoherent to hold both that people have obligations and duties which they recognise and are able to discharge, and that there

is never reason to blame people for their actions. If this were true I could recognise I have a duty to give money to charity and have the ability to do so, but there would be no reason to blame me for not doing so. Or, I could recognise I have an obligation to nurse my dying mother, and have the ability to do so, but there would be no reason to blame me for not doing so. To me, this seems incoherent. The naturalist who thinks these claims are coherent must give us an account of *why* it is right to think I have obligations and duties in these instances. Surely, we abandon the belief that someone has a duty to *x* if we realise that there is no reason to blame him or her for not doing *x*. It appears that we have 'reached bedrock'. If the analytic naturalist were to say 'I know that you think you have a duty to look after your dying mother, and that you know you do, and that you can do it but this does not mean that there is no reason to blame you for not doing so', then I do not know what else we could say to him.

Concerning (ii), the analytic naturalist may question whether it really is incoherent to advocate a view of reasons for action according to which a person might have had no reason to refrain from doing something, even though it is reasonable to blame him for doing it. I think the answer to this question is clear: 'yes – really - it is incoherent'.¹⁸⁶ Moreover, the incoherence is precisely something one ought to recognise in grasping the concept of blame. This is evident in the fact that it is hard to come up with examples to support such a proposal; here is my *attempt*.

My friend is walking home one night when he is attacked. He fights back, leaving his attacker injured. Perhaps the naturalist would claim that there is reason to blame my friend, even though there is no reason for him to refrain from defending himself. Yet, how could this be so? If there were no reason for him to refrain from defending himself, then what *role* would 'blame'

¹⁸⁶ Although some philosophers think it is not so clear. See for example, Svavarsdóttir (2001).

have in this case? Could it be meaningful for the naturalist to blame my friend for his actions? Maybe, if we add things like 'well, he is a black belt in Karate; he has successful averted attacks in the past, he knew his attacker was a weak young girl,' etc. However, if the naturalist attempts to convince us of his position in this way then surely what he is doing is trying to show why the victim *did* have a reason to refrain from defending himself. Moreover, he must think that this is sufficient to secure the judgement of blame. As such, I am unconvinced that the naturalist could present a coherent example which would challenge either (i) or (ii).

Assume then that the naturalist must accept (M*) as true. Therefore, his account is in doubt if he cannot show why his proposed semantic reduction respects (M*). As such, the naturalist must respond directly to the OQA. Simply, the challenge is this: if it is true that analytic naturalism can respect (M*), then why do people have the convictions they do? Specifically, why are competent and reflective speakers of English convinced that they are able to imagine psychologically and linguistically competent people who judge that they have no reason to ϕ in virtue of the fact that ϕ -ing would produce natural property *N*?¹⁸⁷

The challenge to naturalists is to give an account of why we should retain the belief that their account respects (M*) in light of these convictions. They must give an alternative account of why competent and reflective speakers of English have such convictions. There are two ways of doing this.¹⁸⁸

First, the analytic naturalist could argue that the reason that competent and reflective English speakers can imagine what they can is due to the *complexity* of the link between judging

187 Notice that it is not an option for the analytic naturalist to deny that people have such convictions; such a move is empirically false.

188 There is overlap here with the discussion of Ball's paper in chapter II.

that ϕ -ing would produce N and having a reason to ϕ , rather than it not being a conceptual link. The naturalist may cite other conceptual truths in support of such a claim. For example, most people are convinced they can imagine clear-headed beings that fail to recognise that $1 - \sin^2 x$ is equivalent to $\cos^2 x$. The reason is not that there is no conceptual equivalence between them; the reason is that even though they understand the concepts expressed by ' $1 - \sin^2 x$ ' and ' $\cos^2 x$ ', the link is complex and unobvious.¹⁸⁹ In the case of such mathematical examples this seems correct, but it is not obvious that the analytic naturalist can make the equivalent point for morality. Specifically, there seems no *reason* to suspect there is any complexity in the link between moral concepts and natural concepts. For instance, we happily grant that understanding mathematics often requires a long and rigorous training, and that people's intuitions about mathematical claims can be mistaken precisely because we *first* recognise that mathematics is complex. The problem is that we have no such intuitions about morality. In fact, we think the opposite: namely, that people do not need such a training to understand what is (for example) morally right. This suggests that we ought not to draw an analogy between subjects that deal with complex conceptual truths and morality. As such, to make this response convincing, the analytic naturalist is required to show why we ought to take his convictions as correct: that is, as opposed to the convictions of the majority of competent and reflective speakers of English. Presumably, he would argue that it is because *he* grasps the complexity of moral concepts that he can assert

¹⁸⁹ Assuming for the sake of argument that one can talk about conceptual truth in mathematics.

there is such complexity. However, it is unclear what could justify such a claim other than asserting the truth of his position.¹⁹⁰

Rejecting a response in terms of complexity, the analytic naturalist could argue that *systematic confusion* leads English speakers to have such convictions. For example, perhaps they are due to the unrecognised residue of religious belief. However, accepting this does not necessarily demand the conclusion that the OQA ceases to be a challenge. It would only do this if we could show that the confusion generates different intuitions from those which reflect the truth. As a result, for the analytic naturalist to demonstrate that English speakers' intuitions do not count against his proposed semantic reduction, it is not enough for him just to show such intuitions are systematically confused. He will have to demonstrate that the confusion distorts the intuitions people *should* have. To do this he will have to demonstrate that his account is correct.

Nevertheless, even if he manages to demonstrate that the confused intuitions distort the truth, there is another worry. Specifically, it is the intuitions of *reflective* speakers of English that count against the naturalist's proposed reduction. Thus, the naturalist would have to claim that people, *even after reflection*, would remain confused in their intuitions. That is, that there is a systematic confusion which remains, after consideration of the concepts involved. However, it

¹⁹⁰Even if he could provide adequate justification, it still seems reasonable to claim that the OQA would present a challenge if synonymous terms tend to be similarly complex. That is, if we could conclude that typically, terms that mean the same were equally complex, then we could reasonably claim that the naturalist's reduction of the simple term 'good' ought to be simple; we couldn't conclude that people find the naturalist's proposal 'open' because of its complexity. In addition there is evidence that – at least in the moral domain – the naturalist thinks that the terms that mean the same as 'good' are simple. For instance, 'good is pleasure', and 'good is that which we desire to desire'. If this is the case then we can reasonably claim that people do not find questions concerning naturalist's reductions open because they are complex. Hence, the burden of proof lies with the naturalist to demonstrate why it is correct to assert that people's intuitions about his proposed reduction do not challenge his position, as people are ignorant of the complexity of his proposal.

then starts to become hard to understand what entitles the analytic naturalist to claim that his proposed identity is *analytic* and *a priori*. For, even after reflection, people could still fail to spontaneously converge on agreeing to his proposed analysis of moral terms. Hence, whatever cases we could produce demonstrating there is no systematic confusion, the analytic naturalist can insist that we have just not looked at it in the right way - that is, in the way that *he* does. This response then appears to amount to the stipulation that analytic naturalism is correct.

There is one final response available to the naturalist. He could grant all the above, yet point out that the conclusion does not apply to *all* possible reductions, only to the ones that have already been proposed. To this there are number of observations. First, this version of the OQA was set up as a *challenge*, and not as a decisive 'knock down' argument. Second, it is precisely part of the challenge that it asks the naturalist to convince us that a naturalistic reduction *could* respect (M*). That is, given that all proposed naturalistic reductions seem to fail, why think that *any* other account can do better? The analytic naturalist cannot merely respond to this challenge by claiming that they 'just will'. In fact, even the briefest of considerations of our experience suggests that this way of proceeding will fail. The connection between naturalistic judgements and practical reason is elliptical. It is only because we have a certain psychology that a natural property obtaining gives us a reason to act in certain ways. Again, the burden of proof is clearly on the analytic naturalist. How could it ever be *possible* for a semantic reduction of moral terms to natural ones to respect (M*)? I conclude that given that (M*) is true, this version of the OQA presents a genuine challenge to analytic naturalism.

§5. Frank Jackson's Analytic Descriptivism and the OQA.¹⁹¹

Frank Jackson has recently argued for:

... a version of what is often called *definitional* or *analytical* naturalism. However, I will call the doctrine analytical *descriptivism*.¹⁹²

It is the aim of this section to ascertain whether the general remarks I have made in the last two chapters are applicable to the detail and sophistication of Jackson's widely discussed account. I suggest that they are.

Jackson thinks that moral terms are reducible to descriptive terms. He is a cognitivist and not an error theorist and, as such, claims that moral terms have descriptive truth-makers. His reduction is complex and relies on the work of Frank Ramsey and David Lewis.¹⁹³ The motivation behind using this work is the thought that if we fix the meaning of all the terms in ethics by the role or function that they play within ethics, then we can avoid problems other less sophisticated reductions encounter, such as circularity.¹⁹⁴ The details of Jackson's reduction are as follows.

First, Jackson makes a distinction between naïve folk morality and mature folk morality (M). He takes this distinction to be essential to his account. Naïve folk morality is:

¹⁹¹ I will be focusing on Jackson (1998).

¹⁹² Jackson (1998: 113). Emphasis added. He chooses this label to avoid a possible confusion with other questions concerning the relationship between natural and physical sciences and ethics. This need not worry us as in what follows, 'descriptive' can – without any distortion in meaning – be read as 'natural'.

¹⁹³ See Jackson (1998: 140).

¹⁹⁴ For a brief account of the circularity problem, see Kripke (1982).

...the network of moral opinions, intuitions, principles and concepts whose mastery is part and parcel of having a sense of what is right and wrong, and of being able to engage in meaningful debate about what ought to be done.¹⁹⁵

Mature folk morality is:

...where folk morality will end up after it has been exposed to debate and critical reflection (or would end up, should we keep at it consistently and not become extinct too soon).¹⁹⁶

It is \mathbb{M} rather than naïve folk morality which is fundamental to Jackson's account. He starts by imagining \mathbb{M} written out as a huge conjunction of moral propositions written in property name style. For example, 'Killing someone is typically wrong' becomes 'Killing typically has the property of being wrong'. Jackson then replaces each distinct moral property term by a distinct variable to give $M(x_1, x_2, x_3 \dots)$. Then $(\exists x_1) \dots M(x_1, \dots)$ is the Ramsey sentence of \mathbb{M} and $(\exists x_1) \dots (y_1) \dots (M(y_1, \dots) \text{ iff } x_1 = y_1 \ \& \ x_2 = y_2 \ \dots)$ is the modified Ramsey sentence of \mathbb{M} which says that there is a unique realization of \mathbb{M} . Jackson continues:

If moral functionalism is true, \mathbb{M} and the modified Ramsey sentence of \mathbb{M} say the same thing. For that is what holding that the ethical concepts are fixed by their place in the

¹⁹⁵ Jackson (1998: 130).

¹⁹⁶ Jackson (1998: 133). Notice that 'critical reflection' seems to imply further *evaluative* considerations; I return to this below.

network of mature folk morality comes to. Fairness is what fills the fairness role; rightness is what fills the rightness role; and so on. We can now say what it is for some action A to be, say, right, as follows:

(R) A is right iff $(\exists x_1) \dots (A \text{ has } x_r \ \& \ (y_1) \dots (\mathbb{M}(y_1 \dots) \text{ iff } x_1 = y_1 \ \& \ \dots))$.

where ' x_r ' replaced 'being right' in \mathbb{M} . We now have our account of when A is right: it is right just if it has the property that plays the rightness role as specified by the right-hand side of (R), *a property we can be confident is a purely descriptive one*, given the unrestricted, global, a priori supervenience of the ethical on the descriptive. Clearly, the same procedure, with appropriate modifications, will yield an account when A is good, just, fair, bad, and so on. For all ethical predicates, thick or thin, we have an account of their truth – or application – conditions. What is more, we have an account in purely descriptive terms, because the modified Ramsey sentence is obtained by replacing *all* the ethical terms by bound variables.¹⁹⁷

With such roles specified we could investigate which descriptive properties, if any, satisfy them. Jackson gives an example of discovering that the property satisfying the 'rightness' role is maximising expected hedonic value.¹⁹⁸ As such, rightness is identical with maximising expected

¹⁹⁷ Jackson (1998: 140-141). Emphasis mine. A question that I address is how Jackson can be so confident that the rightness-role can't be played by a non-natural property.

¹⁹⁸ Note though that he does not say 'rightness' *means* 'maximising hedonic value'.

hedonic value.¹⁹⁹ Jackson can then give an analytic descriptive account of all evaluative terms: first, specify their role in the appropriate mature folk practice, and second, identify which properties satisfy this role.

Given that Jackson's account is analytic, it seems susceptible to the OQA, a basis form of which would run as follows. One may see that ϕ -ing has the property which would satisfy the 'rightness' role in a network analysed mature folk morality, but is ϕ -ing right? From this we can conclude that Jackson cannot secure the truth of his position. He has two responses to this type of challenge.

Jackson claims that we find such questions open because we ask them from within the context of *naïve* folk morality. However, what matters is the nature of *mature* folk morality because:

... there will, here and now, inevitably be a substantial degree of 'openness' induced by the very fact that the rightness role is currently under negotiation.²⁰⁰

As such, Jackson thinks that people find questions concerning his account open, because they are ignorant of the practice of mature folk morality.

However, there is then a further version of the OQA challenge. That is, even if we knew about the practice of mature folk morality, questions concerning Jackson's account remain open. Jackson responds that at this stage he is entitled to 'dig in his heels' and deny that such a

199 For Jackson's account to avoid the charge of circularity we must read 'hedonic value' qualitatively rather than in evaluative terms.

200 Jackson (1998: 151). Emphasis added.

possibility is relevant to his account.²⁰¹ He attempts to explain why people might have such intuitions, suggesting that it could be due to a 'hangover' from platonistic accounts of meaning. Nevertheless, both of these responses are inadequate and, as such, the OQA remains a genuine challenge to Jackson's account.

As I have presented it, the version of the OQA based on motivation in chapter III would not challenge Jackson's account. This version of the OQA rests on this claim: whereas moral judgements have the capacity to motivate in and of themselves, naturalistic judgements seem to lack such a capacity. Central to this conclusion was my rejection of the belief-desire account. However, Jackson states that such a position is the only plausible account of motivation. Hence, for the OQA to present a genuine challenge based on motivation, I must produce an argument to justify my rejection of the belief-desire account. Without this, Jackson will deny that I am entitled to claim that moral judgements have the capacity to motivate in and of themselves. After all, beliefs, according to him, are 'motivationally inert'. Even given the incompleteness of this challenge, considerations about practical reasons show the OQA to be a genuine argumentative device against Jackson's position.

First, consider Jackson's response that the openness of open questions is due to our ignorance of the practice of mature folk morality. Clearly then he thinks that knowledge of mature folk morality would close such questions. However, I argue that even if knowledge of mature folk did close epistemic possibility, it does not follow that rightness is identical with what plays the rightness role. For Jackson's responses to the OQA to be effective, he has to

201 Jackson (1998: 151).

hold that it is an *a priori* conceptual truth that mature folk are right about 'rightness'. According Jackson, this must be a conceptual truth:

(Q) Rightness is whatever satisfies the 'rightness' role in the network analysed mature folk morality.

Jackson is going to have problems securing the claim that (Q) can be recognised as true purely by grasping the concepts involved. Surely there is nothing in our moral practice that defers to some future time of moral development. Why, for instance, must we think that mature, rather than naïve, folk morality should be definitive of what naïve speakers mean by their moral terms? One way of highlighting this worry is to question how Jackson would respond to the claim that moral practice could mature so that everyone could do what is right, and that this made most people very *unhappy*? To this type of worry he writes:

...a *constraint* on an *acceptable* mature folk morality is that it *should* end up with an account of what rightness is that does not have this consequence.²⁰²

We can now worry about Jackson's use of 'constraint', 'acceptable' and 'should'. It is unclear what he means by them. Perhaps the best way to proceed is to read Jackson as claiming that mature folk are not just future peoples, but are wiser, more reasonable, reflective, and sensitive

202 Jackson (1998:145). Emphasis mine.

peoples.²⁰³ This would make Jackson's point about our lack of knowledge of mature folk practice more plausible. We are more likely to accept that, in fact, in asking if a proposed analysis of 'right' is correct, we are asking whether more rational moral folk would accept such an analysis, rather than asking whether any old future moral folk would accept such an analysis. Reading Jackson's response in these terms makes it more plausible. We find questions concerning his account open, not because his account is problematic, but because we are simply ignorant of what more rational moral folk would accept as a correct analysis.

However, for Jackson to respond in this way he has to deal with the fact that 'rational' is an evaluative term. It seems he has two options; first, he could treat 'rational' in the same way as other evaluative terms, such as 'right'. Second, he could argue that 'rational' is an irreducible evaluative term. I argue that only the second option seems to be available to him.

Recall that to reduce 'right', Jackson claims we have to 'Ramsify' 'mature folk morality'.²⁰⁴ Moreover, we have stated that for his account to be plausible, 'mature folk morality' cannot be merely be the folk morality of any old future generation. It must be the folk morality of a more rational folk. So, if Jackson treats 'rational' as he treats 'right' and other evaluative terms – if he takes the first option – then he will claim that we 'Ramsify' not *our* rationality, but 'mature folk rationality'.²⁰⁵ But what is this? Well, if Jackson is going to be consistent he can't claim that it is the folk rationality of any old future generation, but that it is the folk rationality of... Of what? He will have to answer; it is of a more *rational* generation. But

203 I will use 'rational' to refer to all these evaluative terms.

204 'Ramsify' is my shorthand for the process of reducing terms introduced by Frank Ramsey and discussed in detail above.

205 Assume for the sake of argument that 'mature folk rationality' makes sense.

now Jackson is using precisely what he wants to reduce in his reduction. As such treating 'rational' in the same way as 'right' and other evaluative terms, will lead to circularity.

We can state this as follows. If Jackson wants to reduce 'rational' in the same way that he reduces 'right' then he is committed to something like this:

(RA) ϕ is rational if and only if, ϕ uniquely satisfies the network analysis of 'rational' in mature folk rationality.

So, the property that is rational is the one which is 'picked out' by the 'Ramsified' mature folk rationality. Let's assume, as we did with Jackson's treatment of 'right', that there *will* be one property picked out, and that this property will be descriptive as Jackson insists. Jackson can conclude that 'rational' can be reduced to a descriptive term.

Now, mature rational folk are naïve rational folk who are more rational, but why should 'mature' folk have the role assigned to them in (RA)? How would Jackson respond to the claim that rational practice could mature so that everyone could do what is rational, and that this made most people very *distressed, selfish or unhappy*? Specifically, he cannot respond in the same way as he did with 'right', as to do so would be circular. It would not be illuminating for him to say, "Well, mature folk rationality will not be like this because they are more rational." It turns out, then, that if Jackson wanted to treat 'rational' as he has other evaluative terms, he would be committed to something like (RA), and this turns out to be circular.

We can put this in terms of the OQA. Assume that someone attempts to reduce 'rational' by 'Ramsifying' naïve folk rationality; and that it is a significant and non-trivial question

whether such a reduction is correct. Jackson would presumably claim that this does not challenge his account. We find questions concerning this proposed account open due to the openness of the question 'is x rational?' Moreover, the openness of this question can be accounted for in terms of the question 'is that the most rational way of developing folk rationality from here.' I suggest we would not be satisfied with such a response to the OQA.

The option left to Jackson is to claim that 'rational' is an irreducible evaluative term; 'rational' picks out a non-descriptive property. Concerning such a proposal Yablo writes:

That might look like an impossible position; wouldn't the same supervenience considerations that led Jackson to his descriptivism about ethical apply here too?²⁰⁶

Let's briefly consider what Jackson writes concerning supervenience. He states:

I will start by arguing that the nature of the supervenience of the ethical on the descriptive tells us that ethical properties are descriptive properties in the sense of properties ascribed by language that falls on the descriptive side of the famous is-ought divide.²⁰⁷

What can we say about how the ethical supervenes on the descriptive? Jackson claims that for all worlds if two worlds are exactly alike descriptively then they are exactly alike ethically, and that this is both *a priori* true and necessary. However, Jackson claims that if this were true then

²⁰⁶ Yablo (forthcoming: 7).

²⁰⁷ Jackson (1998: 118-119).

any claim about how things were ethically would be equivalent to some claim about how things are, described in purely descriptive terms.

Next, he specifies that E is a sentence about ethical nature in the following sense:

(a) E is framed in ethical terms and descriptive terms; (b) every world at which E is true has some ethical nature; and (c) for all w and w^* , if E is true at w and false at w^* , then w and w^* differ ethically.²⁰⁸

It follows that in any world that E is true, that world will have some descriptive nature; and there will be a sentence made up of only descriptive terms that gives that nature in full.²⁰⁹ Now, if w_1 and w_2 are the worlds in which E is true, and D_1 and D_2 are the sentences made up of descriptive terms that give the descriptive nature of w_1 and w_2 , then the disjunction of D_1 and D_2 will also be a purely descriptive sentence, call it D . But then:

... E entails and is entailed by D . For every world where E is true is a world where one or other of the D_i is true, so E entails D . Moreover, every world where one or other of the D_i are true is a world where E is true, as otherwise we would have a violation of [supervenience]... Therefore, D entails E . The same line of argument can be applied *mutatis mutandis* to ethical and descriptive predicates and open sentences: for any ethical predicate there is a purely descriptive one that is necessarily co-extensive with it.²¹⁰

208 Jackson (1998: 122).

209 '... ethical nature without descriptive nature is impossible (an evil act, for example, must involve death or pain...).'

210 Jackson (1998: 123).

Jackson concludes that ethical properties are descriptive properties, for it is a consequence of the way that ethical supervenes on the descriptive that any claim about how things are made in ethical vocabulary makes no distinctions among the possibilities that cannot in principle be made in purely descriptive vocabulary.²¹¹

So, to recap, we suggested that Jackson could not claim that 'rational' was reducible in the same way that 'right' 'wrong' and other evaluative terms are. Hence, we claimed that he must be committed to 'rational' being an irreducible evaluative term. However, we then noted that Jackson argued against value being irreducible via supervenience, and the worry was that this argument would apply to 'rational', thus ruling out this second option for Jackson.

We can summarise the possible way Jackson could argue against 'rational' being irreducibly evaluative: the nature of the supervenience of the rational on the descriptive tells us that rational properties are descriptive properties, in the sense of properties ascribed by language that falls on the descriptive side of the famous is-ought divide. This can be expanded. Jackson would claim that for all worlds, if two worlds are exactly alike descriptively then they are exactly alike rationally, and that this is both *a priori* true and necessary. However, if this were true then any claim about how things are rationally would be equivalent to some claim about how things are described in purely descriptive terms.

R is a sentence about rational nature in the following sense. (a) R is framed in rational terms and descriptive terms; (b) every world at which R is true has some rational nature; and (c) for all w and w^* , if R is true at w and false at w^* , then w and w^* differ rationally. It follows that

²¹¹ Jackson (1998: 123).

in any world that R is true, that world will have some descriptive nature; and there will be a sentence made up of only descriptive terms which gives that nature in full. Now, if $w1$ and $w2$ are the worlds in which R is true, and $D1$ and $D2$ are the sentences made up of descriptive terms which gives the descriptive nature of $w1$ and $w2$, then the disjunction of $D1$ and $D2$ will also be a purely descriptive sentence, call it D . But then, R entails and is entailed by D . For every world where R is true is a world where one or other of the D_i is true, so R entails D . Moreover, every world where one or other of the D_i are true is a world where R is true, as otherwise we would have a violation of supervenience. Therefore, D entails R .

We can conclude that rational properties are descriptive properties, for it is a consequence of the way that rational supervenes on the descriptive that any claim about how things are made in the vocabulary of reason makes no distinctions among the possibilities that cannot in principle be made in purely descriptive vocabulary.

If Jackson is committed to this conclusion, given he is committed to the conclusion that ethical properties are descriptive ones, then his position seems to face a challenge. Questions concerning the analytic descriptive account of 'right' are open, and he cannot cite ignorance of mature folk morality accounting for this. For, if he does, he is either committed to circularity or that 'rational' is an irreducible evaluative term. Taking this latter line would in turn mean he has to weaken his original – fundamental – supervenience argument against 'right' being an irreducible evaluative term. Either way, the OQA challenges Jackson's position.

Consider (A):

(A) If ϕ -ing satisfies the 'rightness' role in the network analysed mature folk morality then there is a reason to do that which is ϕ .

Even given Jackson could adopt the claim that 'rational' is an irreducible evaluative term it would not follow that we could see the truth of (A) solely and simply in virtue of grasping the concepts involved. If we cannot reduce 'rational' to descriptive terms, but we can reduce 'right', then (A) is not a conceptual truth.

Thus, given the truth of (M*), Jackson cannot conclude that it is a conceptual truth that rightness is whatever satisfies the 'rightness' role in the network analysed mature folk morality. Therefore, he also cannot merely 'dig in his heels' as a response to the OQA. He has to meet the challenge I have given. This worry is brought out another way. Jackson writes that there could be:

... rather *different mature folk moralities* for different groups in the community; and, to the extent that they differ, the adherents of the different mature folk moralities will *mean* something different by the moral vocabulary because the moral terms of the adherents of the different schemes will be located in significantly different networks.²¹²

If the mature folk moralities of some communities will never converge, then this means that people from these communities could never be in genuine disagreement, but would always be

²¹² Jackson (1998: 137). Emphasis mine.

talking at cross-purposes.²¹³ This is because the meaning of their terms, such as 'rightness' and 'rational', are determined by different mature folk moralities. Therefore, for the people that are part of these communities there would only be apparent disagreements in the same moral language. There can be no such thing as intractable moral disputes between people in different moral communities.²¹⁴

However, because we do not want people who disagree with us to mean something different by their terms, we reject the suggestion that 'reason' and 'right' (and other evaluative terms) are conceptually identical to descriptive properties; as Stephen Yablo writes concerning 'reasonable':

... it seems part of what we have in mind by 'reasonable' that *the door is left open* to the brilliant iconoclast who gets us to see that we have all along been acting contrary to reason. To dismiss such a person as meaning something else by 'reasonable' strikes us as dogmatic.²¹⁵

I agree. What Jackson and other semantic reductionist accounts do not recognize is that we want a moral theory where there are always semantic open questions, one which will not hold that it is an *a priori* conceptual truth that naive folk morality, or mature folk morality, are right about

213 Zangwill (2000) gives the example of Herman Göring and the American prosecutor Robert Jackson at the Nuremberg trials. These men are paradigm cases of people from potential non-convergent mature folk moralities. However, we are convinced that when Robert Jackson claimed that it is wrong to execute Jews and Göring denied this, they were genuinely disagreeing.

214 Jackson then has to give up on one of the central 'platitudes' of folk morality. In fact then his account is, reversionary of our current folk morality.

215 Yablo (forthcoming: 4). Emphasis added to illustrate the 'openness'.

'rightness', 'rationality' etc. Hence, the OQA challenges analytic naturalism. I suggest such a theory, one that avoids postulating moral properties, in chapter IX.

§6. Conclusion

This chapter considered the second suggestion made by DGR concerning the OQA. Their proposal rests on the assumption that there is a conceptual connection between ϕ -ing being right and there being a reason to ϕ . They propose a reason why people are justified in believing that the openness of open questions does not rely on any error or oversight. It is a fact that competent and reflective speakers of English are convinced that they can imagine psychologically and linguistically competent people who judge that they have no reason to ϕ in virtue of the fact that ϕ -ing would produce natural property N . However, if their original assumption were correct, people should not have such convictions. Hence, analytic naturalists cannot secure the truth of their proposed reduction.

I argued that DGR are correct in their original assumption. I demonstrated this by arguing these are conceptual truths: (i) if ϕ -ing is right *and* x is able to tell it is *and* is capable of ϕ -ing, x is blameworthy for not ϕ -ing. (ii) if x is blameworthy for not ϕ -ing then there is a reason to ϕ .

I suggested that, as such, the OQA is a genuine challenge to analytic naturalism. I considered a number of responses that the naturalist could give, that is, ways they could show that the convictions of English speakers have no evidential weight. I found all of these suggestions unconvincing. I concluded that this version of the OQA does present a genuine challenge to analytic naturalism

In the final section, I focused on Jackson's analytic descriptivism and argued that the OQA challenges his account. What the OQA demonstrates is that his position requires irreducible evaluative terms and fails to respect what we deem essential to a moral theory. That is to say, it forces us to adopt an error theory about morality. We think that there is right and wrong in intractable moral disputes; however, we are mistaken. In such disputes, people are talking at cross-purposes – they mean different things by their terms.

I concluded that to think that it is a conceptual truth that mature folk morality is right about 'rightness' and 'rationality' etc., is to be revisionary about folk morality. It is to lapse into dogmatism. In the next two chapters, I discuss non-analytic naturalism which avoids such conclusions. However, I argue that it, too, is susceptible to a version of the OQA.

(V)
NON-ANALYTIC NATURALISM AND THE OQA (I)

Peter Railton and Richard Boyd have put forward distinct naturalist positions which, they claim, are unchallenged by the OQA.²¹⁶ This chapter discusses Railton's account, the next discusses Boyd's. §1 outlines Railton's non-analytic naturalism; §2 considers David Wiggins's unsuccessful version of the OQA; §3 outlines a successful version of the OQA;²¹⁷ §4 draws some conclusions.

§1. Railton's Naturalism.

We can conceive of naturalism methodologically or substantively. As Railton states:

Methodological naturalism holds that philosophy does not possess a distinctive, *a priori* method able to yield substantive truths that, in principle, are not subject to any sort of empirical test. Instead, a methodological naturalist believes that philosophy should proceed *a posteriori*, in tandem with – perhaps as a particularly abstract and general part of – the broadly empirical inquiry carried on in the natural and social sciences.

Substantive naturalism, in contrast, is not in the first instance a view about philosophical methods, but about philosophical conclusions. A substantive naturalist advances a philosophical account of some domain of human language or practice that

²¹⁶ Railton (1986a), (1986b), (1993a), (1993b) and Boyd (1988).

²¹⁷ Wiggins (1993a) and (1993b).

provides an interpretation of its central concepts in terms amenable to empirical inquiry.²¹⁸

These forms of naturalism need not go together. The analytic naturalism discussed in chapters II, III and IV, adopts a substantive claim but not a methodological one.²¹⁹ It states that there is an *a priori* philosophical account of the moral domain that provides an interpretation of moral concepts in terms amenable to empirical inquiry.

Methodological naturalists need not adopt a substantive claim. The broadly empirical inquiry carried on in the natural and social sciences may reveal that we cannot interpret moral concepts in terms amenable to empirical inquiry; such a way of proceeding could favour non-cognitivism, for example.

Railton adopts *both* a methodological and a substantial view even though, importantly, he is a methodological naturalist first. He believes that an *a posteriori* explanatory approach to moral discourse and practice leads to the interpretation of moral concepts as properties that are amenable to empirical inquiry. He writes:

... my naturalist finds [the rejection of substantive naturalism]... unsatisfactory for various reasons, and instead entertains a reduction hypothesis, a synthetic identification of the property of moral value with a complex non-moral property. He does so because he

²¹⁸ Railton (1989: 155-156).

²¹⁹ Wiggins (1993a: 301), calls the combination of substantive naturalism with the rejection of methodological naturalism 'Moore enraging naturalism'.

believes the identification can contribute to our understanding of morality and its place in our world... while preserving important features of the normative role of moral value.²²⁰

Even though Railton is committed to reducing moral value to a complex non-moral property he is not - contrary to what some philosophers have written - scientific.²²¹ He does not think everything in the world, and an individuals place in it, 'belong to' scientific theory.²²² He argues that an *a posteriori* inquiry into how to reduce moral concepts ought not to presuppose any 'test of reality' or 'test of authenticity', and that scientism is 'remarkably crude and hasty'.²²³

Some reductions are eliminative of the reduced phenomena, while some account for it. For example, the failure to reduce caloric fluid, vital force and phlogiston eliminated them, whereas the reduction of water to H₂O reinforces, rather than calls into question, our sense that there really is water. If Railton's project is to succeed, his proposed reduction must vindicate rather than eliminate.²²⁴ He argues that, as such, the property he hopes to identify must be:

... on the one hand, a property to which we have *epistemic and semantic access*, and, on the other, a property which, through identifiable psychological processes, could *engage people motivationally* in the ways characteristic of moral properties.²²⁵

220 Railton (1993b: 317).

221 For example, Wiggins (1993a: 301). calls Railton's account 'crude and scientific'.

222 After all, the naturalist interested in reduction could reject that there is such a thing as 'the scientific method'. See Railton (1993b: 318).

223 Railton (1989: 159).

224 *Ibid.*

225 Railton (1993b: 317). Emphasis added.

The question is, then, whether there is a property that meets such naturalistic and normative requirements.²²⁶

For Railton, the normative requirement amounts to the claim that if something is good then it must be capable of motivating.²²⁷ For example, he writes:

... what is intrinsically valuable for a person must have a connection with what he would find in some degree compelling or attractive, at least if he were rational and aware. It would be an *intolerably alienated* conception of someone's good to imagine that it might fail in any such way to engage him.²²⁸

Initially it is clear that Railton is concerned with the relational 'good for a person' rather than good. The natural property that Railton identifies as good for a person must be particularly 'made for', or 'suited to', whomever it is good for, to such an extent that it is capable of motivating them.²²⁹

Not only is there a normative requirement, but there is also a naturalistic one too. Whatever is good for a person must play an explanatory role, and for it to do so it must meet two conditions:

226 Railton (1989: 163).

227 Although Railton calls this 'internalism', it is distinct from 'internalism' discussed in chapter III. See below.

228 Railton (1986b: 9). Emphasis added.

229 His account of moral value, which I do not discuss, develops from his discussion of non-moral value; the problems of the latter are problems for the former.

1. *independence*: it exists and has certain determinate features independent of whether we think it exists or has those features, independent, even, of whether we have good reason to think this;
2. *feedback*: it is such – and we are such – that we are able to interact with it, and this interaction exerts the relevant sort of shaping influence or control upon our perceptions, thought, and action.²³⁰

I outline Railton's account and then demonstrate how he thinks it can meet the normative and naturalistic requirements.

We best understand Railton's position by starting from his conclusion and concentrating on a number of features: *X* is non-morally good for *A* at time *t* if and only if, *X* would satisfy an objective interest of *A*, the reduction basis of which exists at *t*.^{231 232}

'Subjective interest' includes wants and desires, and should be thought of as akin to taste.²³³ Importantly, subjective interests have relational, dispositional and primary qualities, which Railton calls their 'reductive basis'. These are objective, play an explanatory role, and are 'truth-makers' for subjective interest. However, such a basis is not sufficient to ground non-moral goodness as it does not capture the requisite normativity.²³⁴ We tend to have desires that are irrational, based on incomplete knowledge, or both. Hence, we think that even though we

230 Railton (1986a: 142).

231 For ease, I omit the temporal considerations in the following discussion.

232 Railton (1986a: 143). Railton was, of course, not the first to develop accounts using higher-order desires. Moore (1903) considers them in relation to the OQA and Frankfurt (1971) discusses the importance of second order desires in an account of persons and freewill.

233 *Ibid.* Pg. 142.

234 Railton (1986a: 142).

desire something, it may not be desirable.²³⁵ For this reason Railton characterises objective interest for an individual *A* as what *A* wants and desires, given that *A* has ‘unqualified cognitive and imaginative powers, full factual and nomological information about his physical and psychological constituent, capacities, circumstances, history and so on’ (Railton calls such a character *A*-plus).²³⁶

This granted, Railton’s position is still not an accurate account of non-moral good for *A*, as it concerns good for *A* idealised. Consequently, he claims that the objective interest for *A* is what *A*-plus would want *A* to want, were he in the actual condition and circumstances of *A*.²³⁷ Hence, Railton arrives at an objective interest which, he claims, has the requisite normativity and a reductive basis which grounds the reforming account of ‘non-moral good for *A*’.²³⁸ Railton concludes: *X* is non-morally good for *A* if and only if, *X* would satisfy an objective interest of *A*. However, if Railton’s proposed reduction is to vindicate non-moral good for a person, it must have the normative and naturalistic features highlighted above. I consider these in turn.

If something is good for a person then it must ‘engage people motivationally’.²³⁹ Recall that Railton claimed:

235 For example, consider the drug addict who craves another ‘hit’.

236 Railton (1986a: 142). This is a diverse and complex collection of characteristics and one may wonder whether Railton is really entitled to claim that *A*-plus has the same personality as *A*. For not only must *A*-plus have these features but *A*-plus must also be able to compare all the possibilities that *A*’s life could follow and ‘indwell’ *A*’s actual position so as to advise *A*.

237 ‘We now ask *A*+ to tell us not what he currently wants, but what he would want his non-idealized self *A* to want – or, more generally, to seek – were he to find himself in the actual condition and circumstances of *A*.’ Railton (1986a: 142).

238 The reductive basis here is basically the same as that of subjective interest. The idea of a ‘reforming definition’ can be found in Brandt (1979).

239 Note that this must take place through identifiable psychological processes. Railton (1993b: 317).

... it would be an intolerably alienated conception of someone's good to imagine that it might fail in any such way to engage him.²⁴⁰

Railton gives the following example demonstrating how his account can meet this condition.

Beth is a successful and happy accountant who one day decides to leave her job to become a writer. What Beth doesn't know is that she is, and will always be, at best, an average writer; consequently, she becomes increasingly frustrated and poor. Crucially, Beth-plus would know that Beth would most probably always fail as a writer and would want herself, if she were Beth, to give up writing and return to accountancy, where, she reasons, her non-idealised self will be less frustrated and richer. Therefore, Beth's question, 'Is a writing career good for me?' is answered by considering what Beth-plus would advise Beth to do - precisely what Railton's account suggests.

The important question is how learning Beth-plus's views would affect Beth. After all, Railton introduced the example to justify internalism. He suggests that she would ~~lose~~ the desire to be a writer leaving some contrary desires to emerge.²⁴¹ Railton argues that it is perfectly natural to expect this:

Partly because it is natural to care about whether one is happy and whether one's desires are satisfied. The earlier Beth has every reason to believe that her later self takes these concerns to heart, since the later Beth is contemplating what she would want to pursue

240 Railton (1986b: 21).

241 Railton (1986b: 13-14).

were she actually to relive the intervening years. Moreover, the earlier Beth also has reason to believe that her later self is better situated than she to know what would most satisfy Beth's desires during those years.²⁴²

Thus, the facts that on Railton's account constitute non-moral value will have the requisite normative force; the normative force that facts about non-moral value, truistically, have to possess. The question remains whether Railton's account has the naturalistic feature discussed above. He gives the following example to demonstrate that it does.²⁴³

Lonnie is a traveller who, unbeknownst to him, becomes dehydrated. He feels pretty rotten and homesick and, consequently, longs for something familiar - a drink of milk. Unfortunately, milk is hard to digest and will not help his condition and, in fact, would make him feel physically and psychologically worse. Lonnie-Plus, being fully informed and fully rational, knows that if Lonnie is going to get better he needs to drink a significant amount of clear fluids. It is thus plausible to conclude that what is good for Lonnie is what Lonnie-Plus would want himself to desire if he were Lonnie in these conditions. So, imagine Lonnie given full information unqualified cognitive and imaginative powers, full factual and nomological information about his physical and psychological constituents, capacities, circumstances and history. Then imagine what he would want Lonnie to want. This is very unclear, and it is hard to understand why Lonnie-Plus can be called Lonnie as opposed to someone else, and why he ought to have any authority in making any decisions - I discuss this below. However, assume

242 Railton (1986b: 13).

243 Railton (1986b: 13).

that it makes sense. Then the objective interest (Lonnie-Plus's) with its reductive basis (Lonnie-Plus's dispositions, circumstances and constitution) has these features. He has them, *independently* of whether Lonnie thinks or has good reason to think they exist. Hence, Railton's account satisfies the independence condition. Railton continues the example to demonstrate how it can also meet the feedback condition, and consequently the naturalistic condition.

Lonnie discovers that no milk is available and will not be until the next day. Consequently, he buys a few bottles of water. He drinks these over the next few hours and his psychological and physical health improves. Gradually, when travelling he finds himself drawn to drinking clear liquids rather than milk. This experience *teaches* Lonnie to prefer water to milk when travelling. Railton claims that what this shows is that appeals to degrees of congruence between A's wants and his interests will often help account for facts about how satisfactory he finds his life (that is, the state of his psychological and physical well-being). Furthermore, it demonstrates a way that someone's objective interest could play an explanatory role in the evolution of one's desires; the basic idea being that self-conscious and unselfconscious learning about interests takes place through experience:

In the simplest sorts of cases, trial and error leads to the selective retention of wants that are satisfiable and lead to satisfactory results for the agent.²⁴⁴

It seems then that Railton's account also meets the 'feedback' condition. Lonnie develops the psychological *profile* that he does because of what is good for him. Lonnie can interact with

²⁴⁴ Railton (1986a: 144).

what is good for him, and this interaction exerts the relevant sort of shaping influence or control upon his perception, thought, and action. Not only this but it will also follow that:

...other things equal, individuals will ordinarily be better judges of their own interests than third parties; that knowledge of one's interests will tend to increase with increased experience and general knowledge; that people with similar personal and social characteristics will tend to have similar values; and that there will be greater general consensus upon what is desirable in those areas of life where individuals are most alike in other regards... and where trial-and-error mechanisms can be expected to work well.²⁴⁵

Thus, Railton can claim that the facts in his account that constitute non-moral good can simultaneously play an explanatory role with respect to features of our experience, and can capture the normative role which facts about non-moral value are ordinarily thought to possess. Hence, he claims that the following reduction vindicates rather than eliminates non-moral good: *X* is non-morally good for *A* at time *t* if and only if, *X* would satisfy an objective interest of *A*, the reduction basis of which exists at *t*. The next section discusses, and rejects, a version of the OQA thought to challenge this account.

245 Railton (1986a: 146). Note, one may worry that in this case the causal work is really being done by the reductive basis rather than the fact that something is a non-moral good. Railton claims that there is no need to accept this; to demonstrate why he considers the reduction of water to H₂O. He writes: 'There can be no competition here: the causal work is done by water; the causal work is done by H₂O. Railton (1989: 161). Thus, in the case of the reduction of non-moral good for a person there is no causal primacy, one can say the causal work is done by it, and the causal work is done by the reductive basis: there is 'no competition here'.

§2. David Wiggins's OQA.²⁴⁶

David Wiggins writes that:

Once they are put on to a proper basis, Moore's objections [to naturalism] will reach further (I shall argue) than 'definitional naturalism'.²⁴⁷

The target of Wiggins's 'revived' OQA is Railton's non-analytic naturalism. However, it fails as (a) it begs the question and (b) its scope is too wide.²⁴⁸ First, I outline his argument; second, I develop (a) and (b).

Wiggins partially transposes his argument into a Fregean framework.²⁴⁹ It starts from the assumption that value property V and natural property X are identical.²⁵⁰

(I) $V = X$

(II) A property is natural if a predicate which present it either 'pulls its weight' in an experimental science or is definable in terms of such (otherwise it is non-natural).²⁵¹

246 Wiggins (1993a) and (1993b).

247 Wiggins (1993b: 330).

248 The following discussion is informed by Miller (forthcoming).

249 This is odd as it is of no advantage and complicates matters. Note that he only partially adopts this framework; for example, he talks about properties (something that Frege rejects).

250 The following version is taken from Miller (forthcoming). Presenting the argument in this form brings a clarity which is lacking in Wiggins's presentation.

251 Wiggins (1993b: 332).

- (III) V must have been presented to us under a predicate with the right kind of sense to express a valuational interest.²⁵²
- (IV) X can be presented to us under a predicate with the right kind of sense to express an interest in explanation and prediction (from (II) and the assumption that X is a natural property).²⁵³
- (V) There corresponds to the property V some particular function from objects to truth-values, the V-function.²⁵⁴
- (VI) There corresponds to the property X some particular function from objects to truth-values, the X – function.²⁵⁵
- (VII) The X-function is the V-function (from (I), (V), (VI)).
- (VIII) The X-function projects the valuational interest non-accidentally and faithfully into the future, and across any other cases that could arise (from (VII)).²⁵⁶
- (IX) The valuational interest is not the same as the interest in explanation and prediction.

252 Wiggins (1993b: 332).

253 *Ibid.*

254 *Ibid.*

255 *Ibid.*

256 *Ibid.*

Justification for (IX): Wiggins claims that valuational interest, but not an interest in explanation and prediction, ought to be characterised in terms of *engagement*; he states three ways to understand this.

First, to engage with value involves finding value in an item rather than merely believe it is there.²⁵⁷ But what does this mean? Perhaps it is best to think of this in terms of the claim that value is there to be experienced. But what is it for values to be properties of objects that are there to be experienced? Well, maybe Wiggins means that our concept of value is the concept of a property that is there to be experienced and that this is something that we can show phenomenologically. Evaluative experience presents itself to us as the experience of properties genuinely possessed by the objects that confront us. This argument is to yield the conclusion that objects *seem* to have evaluative properties.²⁵⁸

We can make Wiggins's suggestion more plausible than it may first seem, as it is not obvious that we all do 'find' value in items, that we all experience value as 'there to be experienced'. However, this relies on an assumption that Wiggins need not accept; we need not think of 'there to be experienced' as phenomenologically similar to say, seeing the red of a rose, or finding a table hard. If we do think of it in these terms then it is implausible that we 'find' value in items. However, we need not adopt this assumption. For instance, according to Michael Smith, to talk about finding value in items is to say that:

257 He gives no justification for this claim simply writing that it: '... is more than a verbal manoeuvre'. Wiggins (1993b: 332).

258 I will just consider this first part of the argument. Of course, even given that Wiggins could secure this conclusion, he would have to argue for the is/seems distinction by showing that: '... according to the best systematisation of our platitudes about value, values are dispositions to elicit certain attitudes in us under certain conditions: that is, our concept of value is the concept of such a disposition.' Smith (1993: 242-243) I do not consider this because I judge that Wiggins' will have enough problems with the first stage.

... we would unreflectively describe our experiences as experiences of that property: that is, in apparently representational terms. The crucial question is therefore whether we would unreflectively use [valuational] concepts in the 'predicate position' in describing our evaluative experiences. If we would then it seems that we do have reason to believe that we conceive of values as being there to be experienced... [Moreover] who denies that we would unreflectively describe the experience we have when, say, we witness a wilful murder in apparently representational terms; that is, as the experience of a wrong act?²⁵⁹

We use value concepts in the 'predicate position'; for example, we judge that acts are wrong, and that paintings are beautiful. It is then what Simon Blackburn terms the 'realist-seeming grammar' of our evaluative practice that justifies Wiggins's claim that value is there to be experienced.²⁶⁰

However, there is an obvious problem with proceeding in this way. (IX) requires that we do *not* use natural concepts in the 'predicate position' – which is clearly false. I do not discuss whether Wiggins can produce a phenomenological account that would allow him to further distinguish valuational interest. I leave it that, without further qualification, his first characterisation of engagement is unable to distinguish valuational interest from an interest in prediction and explanation.

Second, Wiggins characterises engagement as:

²⁵⁹ Smith (1993: 243).

²⁶⁰ Blackburn (1985).

... keyed not to the question whether everyone reacts in such and such a way to item *x* *qua* possessed of *V* but to the question whether one is oneself to concur in this reaction.²⁶¹

He believes that the response of finding a value in an item is revisable depending on whether we agree with such a response, whereas judging that something has a natural property is 'keyed into' how everyone reacts. However, consider pleasure, which Wiggins claims is a natural property.²⁶² Finding something pleasurable is surely not 'keyed into' how everyone else would respond. Something being pleasurable depends on whether we find that thing pleasurable. Consequently, this second characterisation also seems incapable of distinguishing valuational interest from an interest in explanation and prediction.

Wiggins's third characterisation states that if engagement connects to feeling then it is 'however indirectly' connected to the will.²⁶³ The scope of this is too wide; it is incapable of excluding an interest in explanation and prediction. Again, consider the judgement that something is pleasurable. Such a judgement is capable of having such an 'indirect' connection with the will, all that is required is that an agent has the appropriate psychology; for example, the desire to do what is pleasurable. Of course, if Wiggins claims that as a matter of conceptual necessity value judgements motivate then his position would be stronger. However, he does not.

261 Wiggins (1993b: 331).

262 '... a natural property is property that pulls its weight within a natural science....On this view, every natural property *either* has some specification or other under which it can figure (and figure essentially) in a natural science *or else* can be defined in terms of properties that have some specification or other under which they figure (and figure essentially) in such a science.' Wiggins (1993b: 330).

263 *Ibid.* Pg. 332.

In conclusion, Wiggins has not successfully distinguished valuation interest and an interest in explanation and prediction. However, for the sake of argument, assume that he can and that consequently he could justify (IX). His argument can then continue:

- (X) It is difficult to see how the X – function could project *both* the interest in explanation and prediction and the valuational interest non-accidentally faithfully into the future and across any other cases that could arise.

Wiggins's justification of (X):

I do not question that, across a fixed number of cases, the function for X could mimic the function for V , and place exactly the same items in its inverse image of the True as the V -function places in its inverse image of the True... [However] in order to mimic the V -function across the open ocean of new cases, surely the X – function has somehow to catch on to the point of the attribution of V . It has to capture and make its own the interest that discerns that value. How else is it to go on indefinitely, generating the same extension as the function for V generates? Can it do this consistently with the claim that the property X is one that can *also* be introduced... by a predicate that pulls its weight in an empirical science or is definable in such terms? The interests seem quite different.²⁶⁴

So, in conclusion:

²⁶⁴ Wiggins (1993b: 332 – 333).

(XI) It is difficult to see how it could be the case that $V = X$.

Even assuming (IX) is true, Wiggins's argument fails on two counts.

(a) *Begging the question*: First, recall from §1 that Railton adopts *both* substantive and methodological naturalism. The property identity claim he proposes does not require *a priori* meaning equivalence. It is via empirical testing that Railton arrives at his reforming account of moral terms. In the context of Wiggins's argument, Railton would conclude $V = X$, only after proceeding *a posteriori* in tandem with the broadly empirical inquiry carried on in the natural and social sciences.

Wiggins claims that the V -function and the X -function cannot 'generate the same extensions non-accidentally faithfully into the future'.²⁶⁵ Railton claims that they can, and there are good empirical grounds for this belief. Wiggins responds that for Railton's account to be correct he must give an *a priori* justification for this. However, this is to *stipulate* that methodological naturalism could never show that substantive naturalism as correct, which is no argument at all. Wiggins begs the question against Railton's position.

We can highlight the point in a different way. Consider Wiggins's claim that an interest in explanation and prediction is different from valuational interest. He thinks that this is problematic for Railton's identity claims. However, it would be strange if this were true, given that Railton explicitly concedes there is such a difference. After all, it is for this reason that he

²⁶⁵ Wiggins (1993b: 332). Notice that this way of talking is very unhelpful. I mean what *does* Wiggins mean by functions 'generating'? As far as possible, we should try to grasp the 'essence' of what he is saying rather than get waylaid by the inappropriate discussion of Fregean functions.

rejects *analytic* naturalism. Wiggins's argument runs: if synonymy is required for property identity, then the difference between valuational interest and an interest in explanation and prediction causes a problem for the naturalist. However, Railton's account denies the antecedent and thus avoids Wiggins's argument - a move the Wiggins simply does not acknowledge. Consider a quotation highlighting this:

Surely if substantive naturalism means what it says when it *identifies* *V* and *X*, then it is committed (and Railton does not demur, it seems) to attempt *very much the same sort of things as an old-fashioned reduction is*.²⁶⁶

This is to miss the central point: Railton's naturalism is *not* attempting 'the same sort of thing as old-fashioned reduction is'. Railton states very clearly that he is a methodological naturalist first, and a substantive naturalist second.

(b) *The scope of Wiggins's argument*: The second problem with Wiggins's argument is that, if successful, it could generate worries about uncontroversial theoretical identity claims. For example, one would have to conclude that there is tension in the claim that water is H₂O, or that salt is NaCl; I take this as a *reductio* of Wiggins's argument. The best way of demonstrating this is by rewriting the argument in terms of an uncontroversial theoretical identity. Thus, assume that water is H₂O.²⁶⁷

²⁶⁶ Wiggins (1993b: 332). *Sic*.

²⁶⁷ This presentation is taken from Miller (forthcoming).

(Ia) $\text{Water} = \text{H}_2\text{O}$

(IIa) A property is natural if some predicate that presents it either pulls its weight in an experimental science or is definable in terms of such. Otherwise it is non-natural.

(IIIa) Water must have been presented to us under a predicate with the right kind of sense to express a practical interest.

(IVa) H_2O can be presented to us under a predicate with the right kind of sense to express an interest in explanation and prediction.

(Va) There corresponds to the property water some particular function from objects to truth-values, the water-function.

(VIa) There corresponds to the property H_2O some particular function from objects to truth-values, the H_2O function.

(VIIa) The H_2O -function from objects to truth-values is the water-function.

(VIIIa) The H_2O -function projects the practical interest non-accidentally faithfully into the future and across any other cases that could arise.

(IXa) The practical interest is not the same as the interest in explanation and prediction.

A brief account of how one might justify (IXa): if I were to have a practical interest in some property, then essentially I want to *do* something with that property. If I were to have a practical interest in gardening, then I would not merely study the right books, I would actively get involved – I would get my hands dirty. A property is practical if this interest picks it out. Such a property is then one that has some function; it is a property with which we do things. Water is a practical property as our practical interest presents it as something we bath in, drink, freeze, etc. Furthermore, to be a practical property requires only one person to have a practical interest in it. It does not require everyone to have such an interest. In addition, if I have a practical interest in water and I have a desire to do what the interest dictates, then I will be motivated to act in certain ways. None of these features are part of an interest in explanation and predication.²⁶⁸

So,

(Xa) It is difficult to see how the H₂O-function could project both the interest in explanation and prediction and the practical interest non-accidentally faithfully into the future and across any other cases that could arise.²⁶⁹

²⁶⁸ Wiggins (1993b: 332).

²⁶⁹ The justification for (Xa) would be the same as for (X); replace 'valuational interest' with 'practical interest'.

In conclusion:

(XIa) It is difficult to see how it could be the case that water = H₂O.

I suggest that this is a *reductio* of Wiggins's argument.²⁷⁰ It also reinforces the 'begging the question' charge. The scientist would grant the claim that 'water' and 'H₂O' have different interests and as such are not synonymous. It is, after all, through a *posteriori* investigation that such theoretical identities are established. However, Wiggins's argument amounts to the assertion that we must base such identity claims on a *priori* meaning equivalence, which is no argument at all. In effect, he stipulates that a *posteriori* investigation could never reveal the nature of water. If Wiggins's account were correct then we would have to question the status of scientific claims in general. Wiggins labels Railton's account as 'crude and scientistic'. In contrast, he not only rejects scientism, but also questions the central claims of science.

Thus, even if Wiggins were correct about the existence of a qualitative difference between valuational interest and an interest in explanation and prediction, his version of the OQA is a failure. He fails to address Railton's position: the combination of methodological and substantive naturalism.

²⁷⁰ Notice that the same form of argument can be run for a whole range of theoretical identities, for example, salt and NaCl.

§3. Railton's Naturalism and the OQA.

Railton writes:

Even if Moore's "open question" argument cannot be deployed directly against an interpretation of discourse about a person's good that does not purport to express analytic truths, a significant critical function may still be served by pressing Moorean questions against such interpretations. For it would be a challenge to any theoretical identification or reforming definition of *P* in terms of *Q* to argue that there is something central to the notion of *P* that does not appear to be captured by *Q*; this would make the question 'I can see that this is *Q*, but is it *P*?' genuinely compelling, not just barely possible.²⁷¹

Railton's account does not rely on an *a priori* meaning equivalence between 'good' and natural terms and hence it is not challenged by Moore's OQA. He thinks that by repeatedly questioning various accounts, we can develop a naturalist account which captures the central notions of non-moral good for a person. However, his mistake lies in where such questions cease to be 'genuinely compelling'. This is because he fails to recognise the role people's ideals play in the normativity of good.

Railton describes personality as:

²⁷¹ Railton (1989: 158).

... a collection of properties that ground dispositions to react in various ways to exposure to certain facts.²⁷²

Railton is an 'internalist';²⁷³ he claims that if something is a non-moral good for someone it must be capable of motivating that person. To reject this is to make good 'intolerably alien'.²⁷⁴

Railton thinks that on gaining full information and rationality, a person's personality *will stay the same*.²⁷⁵

...the idealization holds fixed the individual's non-belief properties, so that the contribution of these features to desire-formation would remain largely the same.²⁷⁶

And,

When we ask [of the idealized agent] how his desires would change upon the impact of further information, we appeal to... [his personality]. We, in effect, hold this [personality] as nearly constant as possible when asking what someone like *him* would come to desire – or, more precisely, would come to want that he pursue were he to assume the place of his original self.²⁷⁷

272 Railton (1986b: 23).

273 Again, this differs from the type of internalism discussed in chapter (III).

274 *Ibid.* Pg. 9.

275 This means Railton can avoid using substantive evaluative judgements, and hence avoid abandoning his reductive programme. Also it means he can retain a continuity between a person and their idealised selves, another essential feature to his account.

276 Railton (1986b: 20).

277 Railton (1986b: 23).

With these points, we can introduce two criticisms of Railton's position. First grant that non-moral good for *A* is what *A-plus* would want *A* to want were he in *A*'s position. Then there seems to be a tension with this and Railton's claim that if something is a non-moral good for *A* it must be capable of motivating *A*.²⁷⁸

The tension arises because Railton thinks that *A* and *A-plus* have the same personality. But if this is the case then how could he cope with a case where *A* hates the personality he has, and hence is not motivated by the decisions it generates. This is a problem as *A-plus* would have the same personality and hence *A* would not be motivated by what advice *A-plus* would give. What *A-plus* would want *A* to want were he in *A*'s position, could then fail to be capable of motivating *A*. It could then fail to meet one of Railton's main criteria for non-moral goodness.

It follows that if this argument were correct, Railton could not 'close' questions concerning his account, even if he rejects his internalism. It is 'open' whether *X* is good for *A*, even if it is what *A-plus* would want *A* to want were he *A*. It suggests that Railton fails to completely naturalise the normativity of non-moral good for a person. I consider this in more detail, attempting initially to reconcile Railton's internalism with his account of non-moral goodness.

First, one may argue that if someone has full information and rationality, then the 'objectionable' parts of his/her personality cease to play a role; something for which (P5) does not account. However, this is problematic. Arguably, something is 'objectionable' in as far as it causes deviation from the non-moral good. Thus, such a response must already assume an account of good.

²⁷⁸ This is Railton's internalism. See for instance, Railton (1986b: 9).

Furthermore, evidence suggests that increasing information does not have a uniform effect on personality. Some people become more compassionate, whilst others become more 'cold' and fatalistic.²⁷⁹ This way of responding by claiming personality changes is difficult. However, this is extraneous, as denying (P5) is not an option for Railton; on his account personality cannot change.²⁸⁰ If it could, then there would be a worry about the continuity between a person and his/her idealised self.²⁸¹

Second response: one may think that it is true that a person could detest his/her fully informed and rational personality, but that what he/she says might still be good for them. Hence, *Railton could argue that the argument rests on the false implicit premise that there is a connection between someone's ideal type of person and what is good for that person.* The problem with such a response is that it puts pressure on internalism (P2). Rejecting the premise Railton is forced to claim that what is intrinsically valuable for a person need not have a connection with what he would find in some degree compelling or attractive, at least if he were rational and aware. In what follows I will outline this in more detail; I start by adapting Railton's example of Beth the accountant.²⁸²

Presume that Beth knows that she is very methodical and un-spontaneous; she is overly thorough, calculates all risks for all actions and tends to lack imagination (this is Beth's personality- something she doesn't much like). Beth-plus knows that Beth will fail as a writer

279 One may be inclined to think that fully informed agents will have the same personality. However, the point is not whether fully informed and rational people would prescribe the same thing but whether what they prescribe is good.

280 Recall that an 'idealised' agent is one with full information and rationality.

281 And problems about circularity.

282 See Railton (1986b), and §1 of this chapter.

and hence wants Beth to give it up.²⁸³ Because she has full information, Beth-plus can recognise *her* personality in a way that is unobvious to Beth. This makes it clear to her why Beth is an excellent bookkeeper, but a lousy writer. According to Railton, the views of Beth-plus will be 'suited to' or 'made for' Beth such that they are 'attractive' and 'compelling' to her; but this does not obviously follow.²⁸⁴ Consider Railton's reasoning again; Beth will be motivated:

... because it is natural to care about whether one is happy and whether one's desires are satisfied.²⁸⁵

The point is that what is important to us, what will be 'compelling' and 'attractive' to us, is not just whether we will be happy. If Beth hates the fact that she is a miser and that this makes her happy, then knowing she will be happy *as a miser* in the future is not going to motivate her. What is important is not whether 'one's desires are satisfied', but whether the desires satisfied are the 'right' ones in relation to the person we want to become. It seems that Railton's internalism is in tension with the response that ideals for people and what is good for them are separable. For instance, one can imagine Beth, on learning Beth-plus's views, saying:

"Well, it is no surprise that *I* say this! Regrettably, I always take the 'safe' option.

However, I do not want to be driven by this personality any more; I promised myself I would act spontaneously to *change myself*. I realise that the only way to do this is by taking risks; but this

283 It is a pressing question why Railton thinks that Beth-plus *knows* Beth's future; however, I do not discuss this.

284 In fact, he says that such a connection is 'psychologically necessary': 'There is no logical contradiction involved in embracing wholeheartedly a desire that one knows one would want not to be effective in one's actual life were one fully informed and rational. The sort of conflict that is basic to the criticism of desires is psychological rather than logical. One might call this conflict "cognitive dissonance" were it merely cognitive.' Railton (1986b: 14 – 15).

285 Railton (1986b: 13).

is fine, even if it means failing as a writer. Such a situation is preferable to staying the way I am!''²⁸⁶

If this description is plausible (and I see no reason to think it is not) then Beth-plus's desires lack the appropriate motivational force. Separating the ideal of a person and what is good for him/her appears to force Railton to reject internalism. We can put this in a general form: assume that A_1 is an ideal *type* of person for A , and A_1 would council A *against* doing X . Then, denying a connection between ideals for people and what is good means it is still possible that X is good for A . However, how could Railton then show that X is 'compelling', or 'attractive' for A such that it is capable of motivating him? Therefore, claiming that there is no connection between ideals for people and what is good for them forces Railton to give up internalism.²⁸⁷ Further problems for Railton's account arise after a few thoughts about the nature of persons.

Assume that what sets people apart from animals is their capability for self-reflection; people have higher-order desires, whereas animals do not.²⁸⁸ They have an ability to 'stand back' from their desires and make various decisions about which 'path' their life will take. We tend not to identify people with their actions and activities.²⁸⁹ Of course, they can be identified in such a way, consider the addict for example. However, in these cases we say that something essential to them as people is lost. Our language reflects this, as we tend to have distinct labels for such characters. In Frankfurt's terminology, they are 'wantons'.²⁹⁰ To have the ability for

286 Perhaps failing a number of times will create in her the spontaneity that she craves?

287 Rosati (1995) makes this point about the nature of persons.

288 This equivalence is controversial; this is how Frankfurt (1971) argues. He uses higher-order desires to characterise 'self-reflection' and 'autonomy'.

289 Roughly, by 'identified' I mean that there is no potential to list the qualities of people. This is different from, for example, tables, dogs and some humans, addicts perhaps.

290 'When a *person* acts, the desire by which he is moved is either the will he wants or a will he wants to be without. When a *wanton* acts, it is neither.' Frankfurt (1971: 14).

self-reflection gives people the ability to invent and re-invent themselves. This involves the ability to consider what *type* of person one wants to become, and what 'path' in life to follow. We do not merely want to be happy, but we desire to experience happiness as a certain type of person.²⁹¹ Arguably, what is central to the concept of person and free will, is the ability for self invention. We can usefully think of the ability to invent oneself as the capacity to 'stand before' a multitude of 'possible selves', and choosing to be guided by one of them. Of course, sometimes we lack the ability to change our personality.²⁹² This, though, is still compatible with hating taking advice from that personality. The point is that we believe we can choose different selves to guide us.

If these thoughts are more or less correct then Railton cannot adopt the second response. Good for a person needs to be 'attractive' and 'compelling' so that it has the capability to motivate. Then it has to be 'suited to' or 'made for' people as *self-inventors*; after all, this is what is essential to personhood. However, Railton explicitly states that to give full information and rationality to an agent does not alter his/her personality. How then could his account be suited to someone as a *person* – as a potential self-inventor? How could it even be capable of motivating that person?

Railton's account does not respect the fact that as persons we have the ability to choose between possible selves. As Rosati writes:

291 Rosati make a useful distinction: '... in determining a person's good, what matters is not simply with whom a person is identical (i.e. the facts about her), but with whom she *identifies*'. Rosati (1995: 60).

292 Frankfurt (1971) and Rosati (1995) both hold a similar view. Again, like many of these points about the nature of personhood, this needs further discussion, but the basic idea is – I think – clear.

[Railton's] definition of 'good for a person' incorporates a person's current motivational system into the very criterion of her good. It thereby treats her as identical with that motivational system for purposes of determining her good.²⁹³

Hence, Railton has a restricted view of autonomy and personhood. The question about what is good for a person and what sort of person to be, are not separable. Rosati again:

In asking what is good, we ask what is worth pursuing or desiring or caring for. But being a particular sort of person involves having characteristic intrinsic motivations, and thereby, a distinctive way of approaching the world, of seeing value in things. Consequently, when we ask what is motivating, we at the same time ask what sort of person to be. We settle simultaneously our questions about who to be, how to desire, and what to seek.²⁹⁴

Because I am a person, an ideal for me (the one I choose to advise me), will dictate what is good for me; the second response is not an option.

Notice that, as a response, Railton cannot simply adapt his account to include substantive evaluative judgements. He claims that it is a *reductive* account of good for a person that he finds most appealing. Of course, it is possible for him to reject reductive accounts, or alternatively internalism (this ability to revise his account is a consequence of his methodological naturalism). However, in

²⁹³ Rosati (1995: 62). Notice that motivation is distinct from action. Arguably, I can be motivated to X, even though I am physically restricted from doing X; motivation ought not to be read as action.

²⁹⁴ Rosati (1995: 62).

doing so he would be abandoning the defining features of his position. In reference to his comments on Moore's OQA opening this section, questions concerning Railton's account are not 'barely possible', but are 'genuinely compelling'.

§4. Conclusion.

Peter Railton puts forward an account that combines methodological and substantive naturalism. It can avoid traditional versions of the OQA because it rejects the need for *a priori* meaning equivalence in establishing property identity. He attempts to give an account of non-moral good for a person, claiming that for it to succeed it must have both a normative and a naturalistic element. Railton concludes that a position based on full information and rationality has these.

David Wiggins has argued that a 'revived' version of the OQA challenges Railton's position. Essentially, he holds that because there is a qualitative difference between valuational interest and an interest in explanation and prediction it seems unlikely that naturalism could be correct. I claimed that such a challenge misses its intended target. Railton *starts* from the claim that there is such difference in interests and hence rejects conceptual analysis. Consequently, he claims to establish his position via *a posteriori* empirical investigation. I concluded that Wiggins's account begs the question against Railton's by assuming that methodological naturalism will never justify substantive naturalism. I also argued that if Wiggins's argument were correct then the truth of orthodox scientific identity claims becomes questionable. I took this as a *reductio* of Wiggins's OQA, and I concluded that it is a complete failure.

However, in the last section I argued that there is a version of the OQA that challenges Railton's account. It turns on the fact that his naturalisation of normativity is incomplete. Questions concerning Railton's account are 'open' because they fail to recognise the role of ideals for people. Possible selves inform our decisions, and the notion of possible selves requires the ability to change personality. This is something that Railton's account rejects. I conclude that the OQA challenges his naturalist position.²⁹⁵

If the naturalist finds this challenge convincing then he could, as discussed above, either give up internalism or reject reductionism and introduce substantive evaluative judgements. In the next chapter, I discuss Richard Boyd's naturalism that takes the former option.²⁹⁶

²⁹⁵ Furthermore, I granted that Railton could talk about full rationality. However it is controversial that one can naturalise full information and full rationality in this way. Thus, even if Railton met the challenge of this chapter he would still have to address this issue. For a good discussion of issues surrounding rationality see Skorupski (1999a) and Scanlon (1998).

²⁹⁶ Boyd (1988).

(VI)
NON-ANALYTIC NATURALISM AND THE OQA (II)

This chapter discusses Richard Boyd's claim that moral terms are natural kind terms that rigidly designate natural kinds.²⁹⁷ I argue that although Boyd's naturalism is resilient to some forms of OQA, it is not resilient to all forms of it.

§1 discusses Boyd's naturalism and §2 considers two versions of OQA, both of which challenge Boyd's account. §3 concludes that non-analytic naturalism based on 'natural kind semantics' fails.

§1. Kripke, Putnam and Boyd's Naturalism.²⁹⁸

The work of Hilary Putnam has provided a foundation for this version of non-analytic naturalism.²⁹⁹ He argues that the 'stereotype' of a natural kind is neither necessary nor sufficient to fix the reference of a natural kind term.³⁰⁰ His 'Twin Earth' thought experiment demonstrates this.³⁰¹ Putnam imagines with the aid of a little science-fiction,

... that somewhere in the Galaxy there is a planet which we shall call Twin Earth. Twin Earth is very much like Earth; in fact, people on Twin Earth even speak *English*... One of the peculiarities of Twin Earth is that the liquid called 'water' is not H₂O but a different

297 Boyd (1988).

298 Kripke (1980), Putnam (1975) and Boyd (1998).

299 Although Putnam and Kripke (1980) both talk about these issues, I will focus on Putnam.

300 By 'stereotype' I mean the features one commonly associates with properties. The 'stereotype' of water would include a liquid which is tasteless, odourless etc. I prefer this to 'nominal essence' as it avoids confusion generated by the fact that Putnam and Kripke argue that natural kinds do not have nominal *essences*.

301 It is possible to generalise this point and claim that 'stereotype' understood in terms of concepts or cognitive content cannot fix the reference of natural kind terms.

liquid whose chemical formula... I shall abbreviate... as XYZ. I shall suppose that XYZ is indistinguishable from water at normal temperatures and pressures..., tastes like water and... quenches thirst like water... that the oceans and lakes and seas of Twin Earth contain XYZ and not water, that it rains XYZ on Twin Earth and not water, etc.³⁰²

In other words, twin-water has the *same* stereotype as water. However, Putnam argues that this does not mean that we Earthians mean by our word 'water' what Twin Earthians mean by theirs. If I were to travel to Twin Earth and pointing to a glass of twin water say 'This is water' then I should have gone wrong - that is - spoken falsely, since, twin water (XYZ) is *not* water (H₂O). This is in contrast to the Twin Earthian who goes right - that is - speaks truly, in pointing to the glass of twin-water and saying 'This is water'. Earthians and Twin Earthians understand the word 'water' differently.³⁰³ The meaning of 'water' tracks a particular kind of liquid, the XYZ kind for the Twin Earthian, and the H₂O kind for Earthian.³⁰⁴ Putnam concludes that it is the environment inhabited rather than the stereotype which determines the meaning of certain terms.^{305 306}

It follows that the 'Twin Earth' thought experiment demonstrates that the referent is the same across all possible worlds. Natural kind terms are, in Kripke's terminology, rigid

302 Putnam (1975: 223).

303 For a clear and useful discussion of this see McCulloch (1995: 167).

304 Kripke (1980: 122), makes the same point about kinds of things.

305 Call these terms 'natural kind terms.'

306 How is the reference of natural kind terms fixed? Kripke suggests it happens via a 'dubbing' with a sample of the natural kind which causally regulate the terms' use. Natural kind terms have their reference grounded by relevant casual 'hook-ups' between speakers and the world, which establishes that natural kind terms' capacity to refer passes throughout the linguistic community.

designators.³⁰⁷ A consequence of this is that natural kind terms flanking an identity sign produce a synthetic identity that, if true, reveals something essential about what is being identified. Thus, although it is an epistemic possibility that water is not H₂O, if it is, then it is metaphysically necessary.

Boyd argues that there is a similar identity between moral properties and natural properties. Even though moral terms and natural terms are not synonymous, if the property identity is true it reveals something essential to moral properties. Thus, the OQA does not challenge his position because there is no *a priori* meaning equivalence between the terms.

Boyd gives an account of how the reference of 'good' is fixed, and how the referential capacity of 'good' spreads throughout the linguistic community:

Roughly, and for nondegenerate cases, a term *t* ['good'] refers to a kind (property, relation, etc) *k* just in case there exist causal mechanisms whose tendency is to bring it about, over time, that what is predicated of the term *t* will be approximately true of *k* ...Such mechanisms will typically include the existence of procedures which are approximately accurate for recognizing members or instances of *k* (at least for easy cases) and which relevantly govern the use of *t*, the social transmission of certain relevantly approximately true beliefs regarding *k*, formulated as claims about *t*...a pattern of deference to experts on *k* with respect to the use of *t*, etc... When relations of this sort

307 Kripke (1980: 48-49).

obtain, we may think of the properties of k as regulating the use of t (via such causal relations).³⁰⁸

§2. OQA and Boyd's Naturalism.

Horgan and Timmons state that if a semantically competent speaker considers a question carefully and judges that the answer to it is obviously 'yes' or obviously 'no', then the question is closed. They claim that a question that is not closed is 'open'.³⁰⁹ Horgan and Timmons's argument against Boyd starts by claiming that (1) & (2) are 'open':

(1) Liquid L is H_2O , but is it water?

(2) Liquid L is water, but is it H_2O ?³¹⁰

This is true. Second, Horgan and Timmons claim that (3) & (4) are 'closed':

(3) *Given* that the use of 'water' by humans is causally regulated by the natural kind H_2O , is liquid L, which is H_2O , water?

(4) *Given* that the use of 'water' by humans is causally regulated by the natural kind H_2O , is liquid L, which is water, H_2O ?³¹¹

308 Boyd (1988: 116). Boyd is then distancing himself from Kripke's discussion of 'baptism' and 'dubbing'; he talks about '... causal mechanisms whose tendency is to bring it about, over time, that what is predicated of the term t will be approximately true of k ...' (1988: 116).

309 Horgan and Timmons (1992: 161).

310 Horgan and Timmons (1992: 161).

Horgan and Timmons argue that on Boyd's account (5) & (6), like (1) & (2), should be 'open':

(5) Entity *e* has natural property *N*, but is it good?

(6) Does entity *e*, which is good, have natural property *N*?³¹²

And (7) & (8), like (3) & (4), should be 'closed':

(7) *Given* that the use of 'good' by humans is causally regulated by natural property *N*, is entity *e*, which has *N*, good?

(8) *Given* that the use of 'good' by humans is causally regulated by natural property *N*, does entity *e*, which is good, have *N*?³¹³

However, Horgan and Timmons suggest that a version of Putnam's 'Twin Earth' thought experiment demonstrates that (7) and (8) are open. Their 'Moral Twin Earth' thought experiment runs as follows:

Suppose that on Earth, 'good' is causally regulated by a functional property whose functional essence is captured by *Tc*, a consequentialist normative theory. Now, consider a

311 Horgan and Timmons (1992: 161).

312 Horgan & Timmons (1992: 155).

313 Horgan and Timmons (1992: 163).

planet 'Moral Twin Earth' which is superficially identical to Earth: water falls from clouds, trains are rarely on time, politicians often avoid giving straight answers to straight questions, and, importantly, Twin Earthians have a moral vocabulary that is orthographically identical to Earthians. That is:

...the terms are used to reason about considerations bearing on Moral Twin Earthling well-being; Moral Twin Earthlings are normally disposed to act in certain ways corresponding to judgements about what is 'good' and 'right'; they normally take considerations about what is 'good' and 'right' to be especially important, even of overriding importance in most cases, in deciding what to do, and so on.³¹⁴

Call these features 'common sense morality'. As such, Earthians would be inclined to interpret Twin Earthians 'good' as meaning the same as theirs.

However, the planets are not identical. Tc does not causally regulate 'Good' on 'Moral Twin Earth', what does is a functional property whose functional essence is captured by a deontological theory Td. Additionally, there is a reliable method of moral inquiry that could demonstrate that Tc and Td played these different roles.

Horgan and Timmons imagine a time before such an inquiry where an Earthling and Twin Earthling disagree over whether something is good. They claim there are two ways of describing such a situation. Despite appearances, the Twin Earthling and the Earthling are not actually disagreeing as they mean different things by 'good'. The environment determines the

³¹⁴ Horgan and Timmons (1992: 164).

meaning of 'good', and they inhabit different environments. Twin Earthling and the Earthling are best thought of as talking at 'cross-purposes'. If this is the correct way to proceed then the intuitions match those in Putnam's original thought experiment. Alternatively, common sense morality determines the meaning of 'good', and there is genuine disagreement.

Hogan & Timmons propose that the second description is the most natural. Learning that Td, rather than Tc, causally regulates 'good', does not invite the conclusion that it differs in meaning. Consequently – contrary to what Boyd thinks – our semantic intuitions do not support the claim that a natural kind causally regulates the use of 'good'.

Horgan and Timmons argue that their conclusion is perfectly general across all non-analytic naturalistic accounts:

For any [version of non-analytic naturalism]... according to which (i) moral terms bear some relation R to certain natural properties that collectively satisfy some specific normative moral theory T, and (ii) moral terms supposedly *refer* to the natural properties to which they bear this relation R, it should be possible to construct a Moral Twin Earth scenario suitably analogous to the one constructed above – i.e., a scenario in which twin-moral terms bear the same relation R to certain natural properties that collectively satisfy some specific normative theory T', incompatible with T. The above reasoning [against Boyd's account] should apply, *mutatis mutandis*, against the envisioned alternative version of [nonanalytic naturalism].³¹⁵

315 Horgan and Timmons (1992: 167)

If non-analytic naturalists want to resist Horgan and Timmons's OQA, then they have four options.³¹⁶ First, they can deny that Horgan and Timmons's account describes a genuine possibility, and so hold that our intuitions have no evidential weight. This response is too strong. It would require defending the conceptual impossibility of such thought experiments. Given the 'Moral Twin Earth' scenario is intelligible, this would incur an overwhelming burden of proof.

Second, they can argue that although (7) & (8) are 'open', there is no guarantee that they will stay this way; a natural property may end up 'closing' them. Apart from an unrealistic optimism, this response fails because it disregards the fact that Horgan and Timmons's OQA is a general argumentative strategy. Consequently, such a reply would have to demonstrate the inadequacy of OQA before making such claims about the future.

Third, they can claim that people's intuitions about 'Moral Twin Earth' are simply mistaken and hence do not undermine causal semantic naturalism:

After all, brute intuitions don't settle the matter by themselves; they can be overruled by a semantical theory that best accommodates all relevant data, including data about the workings of moral discourse and inquiry. Once all relevant data are considered, the new waver might argue, Boydian semantics has a lot going for it and recalcitrant intuitions should be dismissed.³¹⁷

³¹⁶ They could also adopt a form of relativism. However, I take Horgan and Timmons's response to this as definitive: [Relativism] will be very unattractive to new wave meta-ethical naturalists like Boyd. These philosophers espouse *moral realism*, an adamantly non-relativist position. And in any case, relativism is the last defensive ditch against Moral Twin Earth. If ethical naturalism gets forced into that ditch, there it deserves to die. Horgan and Timmons (1992: 170).

³¹⁷ Horgan and Timmons (1992: 168).

Such semantic intuitions could be mistaken; however, the burden of proof would be great, as:

First, we are talking about not just any old intuitively strong beliefs involving morality, but intuitive beliefs about *meaning*, surely an important kind of data when one is considering theoretical accounts of meaning, not so easily dismissed. Second, even allowing that such intuitions could be overridden by an otherwise plausible semantics of moral discourse, nevertheless, *ceteris paribus*, a semantical account that respects meaning-intuitions of competent language users is to be preferred over one that does not.³¹⁸

The fourth response focuses on the work of Laurence, Margolis, and Dawson – I will argue this, too, fails.³¹⁹ They claim that Horgan and Timmons's argument is 'a cheat' as it relies on a commonly known chemical identity. They write that the answers to (3) & (4) do seem obvious to ordinary English speakers. Yet this is simply because it is such a familiar everyday fact that water is H₂O. The answers to these questions are, obviously, 'yes' simply on the ground that *everyone knows* that water is H₂O.³²⁰

Horgan and Timmons's argument then does not rely on semantic intuitions but on how much chemistry people know. The correct response to the 'Moral Twin Earth' thought experiment is to deny that (3) & (4) are 'closed'. There is then no 'open/closed' distinction and

318 Horgan and Timmons (1992: 169-170).

319 Laurence, Margolis and Dawson: <http://www.shef.ac.uk/~phil/staff/laurence/papers/MRTE.html>.

320 Laurence, Margolis and Dawson: <http://www.shef.ac.uk/~phil/staff/laurence/papers/MRTE.html>: 12.

hence no argument against Boyd.³²¹ To make this clear, Laurence, Margolic, and Dawson claim that (7) & (8) are not analogous to (3) & (4), but are analogous to questions where the chemical identity is not so familiar, like (9) & (10):

(9) Given that the use of 'saccharin' by humans is causally regulated by $C_7H_5O_3NS$, is substance S, which is $C_7H_5O_3NS$, saccharin?

(10) Given that the use of 'saccharin' by humans is causally regulated by $C_7H_5O_3NS$, is substance S, which is saccharin, $C_7H_5O_3NS$.³²²

They claim that because (9) & (10) are 'open', but 'saccharin' does rigidly designate the natural kind $C_7H_5O_3NS$, our semantic intuitions lack evidential weight. Specifically, we can conclude that (7) & (8) are 'open', but that 'good' rigidly designates N. However, I shall now argue that they are wrong; (9) & (10) are 'closed', and (7) & (8) are analogous with (3) & (4).

Somewhere in the Galaxy there is a planet which we shall call Twin Earth.³²³ Twin Earth is very much like Earth; in fact, people on Twin Earth even speak English. One of the peculiarities of Twin Earth is that the substance called 'saccharin' is not $C_7H_5O_3NS$ but a different substance whose chemical formula I shall abbreviate as XYZ. I shall suppose that XYZ is indistinguishable from saccharin at normal temperatures and pressures, tastes like saccharin and sweetens coffee like saccharin. The shops, homes and teas of Twin Earthians' contain XYZ.

321 'Semantic open/closed distinction' refers to whether, based purely on semantic intuitions (rather than epistemic state), one would answer a question obviously 'yes' or obviously 'no' (closed), or whether one would not (open).

322 Laurence, Margolis and Dawson: <http://www.shef.ac.uk/~phil/staff/laurence/papers/MRTE.html>: 12.

323 This is adapted from Putnam (1975: 223).

In other words, twin- saccharin has the *same* stereotype as saccharin. However, this does not mean that we Earthians mean by our word 'saccharin' the same thing as Twin Earthians. If I were to travel to Twin Earth and pointing to a dispenser of twin saccharin say 'This is saccharin' then I should have gone wrong, that is, spoken falsely, since twin saccharin (XYZ) is *not* saccharin ($C_7H_5O_3NS$). This is in contrast to the Twin Earthian who goes right, that is, speaks truly, in pointing to the dispenser of twin saccharin and saying 'This is saccharin'. Earthians and Twin Earthians understand the word 'saccharin' differently. The meaning of 'saccharin' tracks a particular kind of substance, the XYZ kind for the Twin Earthian, and the $C_7H_5O_3NS$ kind for Earthian.³²⁴ It is the environment inhabited rather than stereotype which determines the meaning of the terms. If this is true, as it seems to me to be, then (9) & (10) are closed and Laurance, Margolis and Dawson are mistaken. That is, the 'revived' OQA does not depend on how much people know about chemistry. It assumes that we know the referent of 'good' and then considers our semantic intuitions.

None of the four considered responses to Horgan and Timmons's argument is successful. However, things are worse for the non-analytic naturalist. There is a version of the OQA that does not rely on specific intuitions about particular thought experiments.

What makes natural kind reductions such as Boyd's possible is the claim that certain moral terms are rigid designators. A term being a rigid designator relies on certain *referential intentions*. Hence, if there is an asymmetry between the referential intentions associated with rigid designators and those associated with moral terms then this would count against Boyd's position. Hence, to judge the success of Boyd's claims we must clarify the nature of the

324 Kripke (1980: 122), makes the same point about kinds of things.

referential intentions associated with rigid designators. To this end, I consider the rigid designator, and natural kind term, 'water'.

Using the term 'water', we intend to refer to a property that has a particular role. In telling someone that a glass is full of water, we communicate that the liquid in the glass will have certain effects on him or her. It will *cause* him or her to see a liquid that is transparent, quench his/her thirst, etc. It will cause those characteristics typically associated with water.

Furthermore, 'water' is used to 'pick out' what is *actually* causally responsible for these characteristics. We are not, for example, concerned with the property causally responsible in another possible world. This reflects our belief that in using 'water' we intend to refer to that which we encounter in our everyday experience. We are interested in referring to whatever it is that causally impinges on our actual lives.

Finally, by using 'water', we 'pick out' a property that may lack some or all of the characteristics typically associated with water. If water were in fact green, then we would not conclude that water ceased to exist; the conclusion would be that water had always been green. This holds for all the characteristics associated with water: water may turn out to be a solid, boil at hundred and ten degrees, taste of Cola, etc. We can put these three points as follows:

(W) In using the term 'water' speakers intend to refer to whatever property was, in the actual world, causally responsible for water's ordinary characteristics and causal role.

I suggest this can be generalised for all natural kind terms that are rigid designators:

(N) Speakers intended to refer to whatever property was, in the actual world, causally responsible for the kind's ordinary characteristics and causal role.³²⁵

If non-analytic naturalism were correct then a similar statement ought to be available; for example:

(G) In using the term 'good' speakers intend to refer whatever property was, in the actual world, causally responsible for the characteristics, including the causal role, common sense morality characterises for good.

(G) is not something we immediately recognise as true. That is, we do not straight away see that these are our referential intentions when applying 'good'. This suggests that 'good' is not a rigid designator. However, before we can confirm this suspicion, further discussion is required concerning our referential intentions.

Saying that something is 'water' communicates that it will have certain causal effects: if someone drinks it, it will quench his or her thirst; if it is taken to zero degrees centigrade it will solidify, etc. However, communicating to others that certain facts have specific causal roles is not part of our referential intentions with moral terms. When we say that something is wrong, we are trying to get people to disapprove or feel a certain way; we are trying to get people to recognise it is something that there is reason not to do.

³²⁵ See Gampel (1996: 196) also has this formulation.

The non-analytic naturalist may respond by claiming that there are in fact certain causal roles that people intend to refer to with their moral terms. However, this seems highly improbable. While there is a consensus concerning the causal role of temperature, water, salt, etc., there is no consensus concerning the causal role of good, wrong, right etc. Such asymmetry is a problem for the non-analytic naturalist. It seems unlikely, then, that there is a causal role that we all *intend* to refer to when using moral terms. In fact, I suggest that if there were such a role it would itself be moral in character. For example, good people do good deeds; injustice causes harm. This makes trouble for a naturalist who seeks to provide a reductive account of the relevant causal roles and moral properties.

Could it be the case that the original users of the terms 'good', 'bad', 'right', 'wrong', etc. picked out things which did not have the ethical properties in question? If we think the answer is 'yes' then it appears that another of the referential intentions identified does not apply to moral terms, for it is not the case that we think the original creators of 'hot', 'water', 'salt', etc. could have picked out things that did not have the natural properties in question. As Kripke has argued, there is in such cases an original 'dubbing', or 'baptism', of such terms with actual instances of the property. It is because of this that a 'referential link' can be 'spread' through the linguistic community.³²⁶ I suggest that the answer to this first question is 'yes'; the original use of 'good', 'right' etc. could have picked out things that did not have the ethical properties in question. It seems that anyone could have been wholly mistaken in his or her ethical statements. This does not mean that there is no way of knowing what is good - just that no one knows yet.

³²⁶ Kripke (1986:133-139).

Although improbable, this is possible, and hence suggests that there is another difference between the referential intentions with rigid designators and moral terms.

The final referential intention highlighted also fails to apply to moral terms. It is true that when using the term 'water' we accept that all the characteristics we associate with the property water may turn out to be different. However, it is not true that in using the term 'good' we accept that all the associated characteristics may turn out to be different. If what I have argued for in earlier chapters were correct, then the ability to motivate and the connection to practical reason are fundamental to our understanding of moral terms. Therefore, in using moral terms, these features must be present if we are to refer to various moral properties. However, the point need not rest on the truth of these particular claims. The question is whether in using moral terms we intend to refer to properties that may lack all of those characteristics we associate with moral properties. It seems to me that in such cases we would not conclude that our judgement about good was mistaken, but that we cease to be referring to good. Importantly, this is not to claim that all the ordinary characteristics we associate with good are correct; I am not a perfect judge of what is good, and can imagine that some ordinary criteria are wrong. The point is that what leaves this 'openness' to change is my awareness of alternatives whereas, with the criteria for natural kinds:

...my Agnosticism...is due to my knowing they are crude approximation methods, and that there are stricter (causal) ones which trump them.³²⁷

327 Gampel (1997a: 159).

These observations highlight that the burden of proof lies with anyone who wants to claim that moral terms function like terms denoting (natural) kinds. So, there is a way of challenging non-analytic naturalism that relies on everyday intuitions, rather than those based on 'Moral Twin Earth'.

§3. Conclusion.

Nonanalytic naturalism can respond to the OQA discussed in chapters III and IV, as it rejects the claim that property identity requires synonymy. Boyd argues that there is a synthetic identity between moral properties and natural properties that, if true, reveals something about the essence of moral properties.

I argued that there are two related versions of the OQA that challenge non-analytic naturalism. The first states that if Boyd were correct and moral terms rigidly designate natural kinds then a suitably altered version of Putnam's 'Twin Earth' thought experiment ought to support this. However, the thought experiment does not, and hence Boyd cannot secure the metaphysical necessity required for his account.

I claimed that, although a genuine challenge, the OQA would be stronger if it did not depend on a specific thought experiment. Therefore, I suggested another way of capturing the same open question intuitions that does not rely on claims about 'Moral Twin Earth'. Reflection on one's semantic practice reveals that there is a significant difference between the referential intentions with moral terms and the referential intentions with natural kind terms. This difference counts in favour of treating the OQA as a genuine argumentative device.

I conclude that the OQA not only challenges analytic naturalism but also challenges non-analytic naturalism. In the next chapter, I discuss supernaturalism as a way of clarifying the relation between open questions and epistemic possibility.

(VII)

ROBERT ADAMS' SUPERNATURALISM

AND THE OQA

In this chapter, I discuss Robert Adams' claim that certain moral properties are supernatural.³²⁸ In §1 I outline Adams' version of the OQA; in §2 I argue that the 'openness' of open questions is not due to epistemic possibility, that Adams' version of the OQA fails and that his account is challenged by the OQA; in §3 I draw some conclusions.

§1. Good as a Supernatural Property.

Adams claims that Good is God. God is the exemplar of good; that is, things are Good in as much as they resemble God. However, this does not mean that 'Good' is synonymous with 'resembles God'. Rejecting the idea that property identity requires *a priori* meaning equivalence, Adams adopts a semantic account similar to Richard Boyd's.³²⁹ To understand 'Good' is to grasp the characteristics of Good, and these fix the reference of 'Good'. Once this reference is discovered *a posteriori* it can be claimed that Good, as a matter of metaphysical necessity, is identical with that referent. Adams' account differs from Boyd's by arguing that Good is not a natural kind.³³⁰ Even so, Adams writes:

328 Adams (1979) and (1999).

329 Boyd (1988). See chapter VI.

330 Adams writes: What is given by the meaning [of 'good']... is a role that the nature is to play. If there is a single candidate that best fills the role, that will be the nature of the thing. In the case of a natural kind, arguably, the role its nature is assigned by our language is that of accounting causally for the observable common properties of identified samples. The role that the meaning of 'good' picks out for the nature of the good will be rather different. Adams (1999: 16).

... the meaning of the word 'good' may be related to the nature of the good in something like the way that has been proposed for natural kinds.³³¹

Adams' project has two stages. First, it specifies the nature of Good and second, he considers whether there is a property with this character. However, Adams claims that to try to specify the nature of good in all contexts is too hard. Hence, he considers the idea of Good understood as *excellence*. His first stage, then, is to specify the characteristics of excellence.

States of affairs can have instrumental value and can be desirable for their own sake, but they lack a certain sort of unity. Consequently, Adams argues excellence is not one of their properties. He claims that this is something we can learn by reflecting on language. Beauty is an example of excellence and we say persons, physical objects, deeds and lives and some kinds of abstractions, can be beautiful, but states of affairs cannot.³³² Excellence is not, then, a property of states of affairs.

Even so, Adams argues that excellence is best thought of as a property. He suggests that grammar supports the claim that excellence is an objective property; that is, we talk as if excellence were a property that existed independently of people's judgements.³³³ In our inquiries about what is, or is not, good we generally think and speak on the assumption that we could be mistaken. The second characterisation of excellence is, then, that it is an objective property.

Clearly, people pursue what they judge to be excellent. Adams therefore considers the nature of this pursuit. As evidently, people pursue things which are not excellent, such as

331 Adams (1999: 16).

332 Plato also argued that beauty (καλός) is excellence. See Plato's Symposium (210E).

333 Adams (1999: 18).

money. To clarify the nature of the pursuit of excellence, Adams introduces the Greek idea of *Eros*. Although it has the most sexual connotations of any of the main Greek words for love, *Eros* is not merely a physical drive but also prizes the object it pursues as intrinsically valuable. Furthermore, *Eros* does not pursue without reflection; we admire the object of *Eros* and due to this admiration, we pursue the object. Adams claims that this is similar to how one understands the pursuit of excellence. Rather than being physically driven to pursue excellence, one first admires what is excellent as intrinsically valuable, and then pursues it. However, Adams asks:

... don't we admit that, in our folly, we often love, with admiration and desire, things that are not... [excellent], and fail to love things that are?³³⁴

This however is compatible with Adams' account.³³⁵ He attempts to outline the nature of Good understood as excellence that fixes the reference of 'Good', rather than give a definition. Adams thinks it supports his position that things could exhibit the characteristics of excellence yet not be excellence. Hence, the admission in the quotation does not threaten Adams' account. However, a stronger claim is also a possibility on his account, namely, that things *always* or *typically* exhibit all the characteristics of excellence yet are not excellent. Adams wants to reject this stronger claim since, if it were accepted, a property typically belonging to things judged to be bad could be excellent. He thinks that, typically, we can know whether something is good or not, and consequently we cannot usually or typically be mistaken about goodness. Excellence,

334 Adams (1999: 19).

335 Just as, arguably, something can have all the features commonly associated with water, and yet not be water.

to some extent, is a property that one can *recognise* in things.³³⁶ Thus, for Adams, excellence must be an object of recognition; it must be the case that excellence is normally recognised as excellence. He claims this does not commit him to conclude that whether or not something is excellent depends on whether it seems to be excellent. He still maintains that we can be wrong about what is excellent – but not on a regular basis.

From this discussion, Adams draws the conclusion that:

The thesis... that seems to me most clearly correct is that to the extent that anything is good, in the sense of 'excellent', it is *good* for us to love it, admire it, and want to be related to it, whether we do in fact or not.³³⁷

The question remains as to what property exhibits these characteristics. Is there a property that is good to love, admire, and want to be related to, whether we actually do or not? Adams thinks that there is - the property of resembling God. If it is good to love or admire a thing then this is because that thing resembles God. As far as something resembles God, we can conclude that if that thing were loved and admired, then it would be good. Consequently, via a modern semantic account, Adams demonstrates that the proposition that good is resemblance to God is a synthetic yet necessary identity.³³⁸ However, for Adams' project to be successful, he

336 Adams (1999: 20). Surely though this admission is controversial. That is, it looks as if Adams is admitting what he takes to be the 'real essence' of good into the 'nominal essence.' I discuss this in more detail below.

337 Adams (1999: 20).

338 Furthermore, resemblance to God is something which things have or lack objectively. Hence, Adams labels his position moral realism.

must demonstrate why a supernatural, rather than a natural, property best satisfies the features of excellence.³³⁹ He thinks he can secure this by rejecting Boyd's account.

Adams claims that Boyd's position has a problematic consequence. One could discover the natural kind tracked by 'good' and hence questions concerning Boyd's account would be 'closed'. He labels the reason that such a consequence is problematic the 'critical stance'; what he calls the 'truth behind' the 'open question' argument.³⁴⁰ Adams writes:

The stance amounts to at least this. For any natural, empirically identifiable property or type of action that we or others may regard as good or bad, right or wrong, we are committed to leave it *always* open in principle to raise evaluative or normative questions by asking whether that property or action-type is *really* good or right, or to issue an evaluative or normative challenge by denying that it is *really* good or right.³⁴¹

According to Adams, we ought to reject Boyd's account because it does not respect the critical stance.³⁴² He thinks the following is possible on Boyd's account: society will at some point discover the nature of good and once this has occurred we will have to concede that one cannot significantly and meaningfully question this discovery. At this point, Boyd's account becomes incompatible with the critical stance. It is committed to closing the epistemic possibility concerning the nature of good. In contrast, Adams writes about his own view:

339 Boyd (1988).

340 Adams (1999: 78).

341 Adams (1999: 78).

342 A similar thought can be found in Jackson: in his words, immature folk morality respects the critical stance, whereas mature folk morality does not. Jackson (1998).

A realist view of the Good as transcendent favours a particular version of the critical stance. It is a stance that stands ready to question or challenge *any* human view about what the Good is, and any empirical test of values. At the same time it is not a stance of arbitrariness or mere self-expressiveness... It reaches out towards an objective standard that is actually glimpsed, *though never fully or infallibly*.³⁴³

Hence, Adams argues that his account can always respect the critical stance; it is because the nature of God is beyond human knowledge that in principle one can always raise evaluative or normative questions by asking whether the proposed property or action-type is really good or right. In conclusion, 'good' understood as excellence is identical with resemblance to God, and this, unlike naturalism, is compatible with 'a very important feature of ethics' - the critical stance.³⁴⁴

§2. Analysis of Adams' Account of Good.

In this section, I first argue that Adams' objection to Boyd's naturalism is mistaken. Second, I argue that because the OQA is not about epistemic possibility, it challenges Adams' position.

343 Adams (1999: 82). Emphasis added. Notice that in chapter IX I suggest that we can account for this semantic practice without having to postulate a transcendent Good; our semantic practice with moral terms can capture the evidence transcendence usually thought to be evidence for moral realism.

344 Adams (1999: 78).

Adams argues that Boyd's account is incompatible with the 'critical stance'.³⁴⁵ Leaving aside that the OQA is not about epistemic possibility, is he correct?³⁴⁶ Boyd thinks that moral terms are natural kind terms that rigidly designate natural kinds. Assume then that we know that good is pleasure (there are further questions concerning the relationship between knowledge and certainty which I do not discuss) we then *know* that 'good' rigidly designates pleasure; we know 'good' tracks pleasure across all possible worlds and there are no worlds in which 'good' refers to something else. Consequently, on Boyd's account, *if* we know that good is pleasure, then it is not metaphysically possible that good is something else.³⁴⁷ However, does this mean Boyd is committed to the possibility that we *can* have such knowledge? No, he argues that we could be wrong in our beliefs about what we think we know, even if we judge these beliefs justified. Therefore, Adams is mistaken in thinking that Boyd's account will necessarily close 'epistemic' possibility. Hence he is unable to establish that Boyd's account fails to respect the 'critical stance'.

My second argument is that the OQA challenges Adams' position. Given that Adams' account identifies good with the *unknowable* God, I conclude that the OQA has nothing to do with epistemic possibility.³⁴⁸

Horgan and Timmons claim that there is a version of the OQA that applies to Boyd's nonanalytic naturalism.³⁴⁹ They close epistemic possibility and then consider semantic intuitions. That is, they propose a test that runs as follows; first, grant that we know certain facts,

345 Adams (1999: 78).

346 Adams (1999: 81).

347 This is explained in more detail in the last chapter.

348 See chapter VI.

349 Horgan and Timmons (1992). Their argument is a way of formalising the point I made above concerning epistemic and metaphysical possibility.

then consider questions concerning certain synthetic identities 'closed'. Horgan and Timmons claim that we ought to find such questions 'closed'. They, for instance, argue that the following questions are 'closed':

(1) Given that the use of 'water' by humans is causally regulated by the natural kind H_2O ,
is liquid L, which is H_2O , water?

(2) *Given* that the use of 'water' by humans is causally regulated by the natural kind H_2O ,
is liquid L, which is water, H_2O ?³⁵⁰

Horgan and Timmons argue that if we consider analogous questions on the *Boydian* account, then our semantic intuitions concerning it are different. *Boyd's* account fails the 'test' for synthetic yet necessary identities, and hence Horgan and Timmons conclude that in the moral case there cannot be these types of identities. They conclude that *Boyd's* account relies on an implausible semantic framework.

Given that Adams employs the same semantic framework as *Boyd's* account, we can assume that Horgan and Timmons's OQA will challenge his account. Such an argument would proceed as follows. Initially, we have to 'close' epistemic possibility concerning good's identity with resemblance to God, and then we consider our semantic intuitions concerning this identity. If these intuitions mimic those concerning (1) & (2) then Adams' account differs from *Boyd's*

³⁵⁰ Horgan and Timmons (1992: 162).

and hence could possibly avoid the OQA. Initially, to see if this is plausible we must find questions equivalent to (1) & (2):

(3) *Given* that the use of 'good' by humans is regulated by the property of resembling God, is L, which resembles God, good?

(4) *Given* that the use of 'good' by humans is regulated by the property of resembling God, does L, which is good, resemble God?

The key question is whether semantically competent agents would judge these questions 'closed'. If they would, then we can conclude that the type of question being asked in (3) & (4) is similar to the type of question being asked in (1) & (2). This means that the OQA does not challenge Adams' account in the same way that it challenges Boyd's position. However, I claim that the opposite is true. That is, our response to (3) & (4) is fundamentally different from our response to (1) & (2). However, to show this we need a further argument.

Horgan and Timmons use Putnam's 'Twin Earth' thought experiment to demonstrate that the answers to (1) & (2) are obviously 'yes' or 'no'. Hence, Adams may be able to adapt Putnam's thought experiment to demonstrate that the answers to (3) & (4) are obviously 'yes' or 'no'. Of course, if the opposite were true – that the answers to (3) & (4) are not obvious – then this would establish that Horgan and Timmons's OQA presents a genuine challenge to Adams' account. That is, we were right in the first place; (3) & (4) are 'open' questions.

Suppose that 'Twin Earth' and Earth are alike, apart from Earth having a God. Importantly, Earthlings and Twin Earthlings use 'good' in identical ways, such that Earthlings would be convinced that Twin Earthlings and themselves mean the same by 'good'.³⁵¹ This identity includes the type of action guiding role discussed in earlier chapters.³⁵² Consider then a case where an Earthling finds himself on 'Twin Earth' and also finds himself engaged in a moral disagreement concerning whether something is good. If the analogy with Putnam's thought experiment holds, then our semantic intuitions ought to be that the Earthling and Twin Earthling are not disagreeing. We must conclude that they are talking at cross-purposes.

However, this description is not the most natural. Given that the action-guiding role of 'good' is identical across Earth and 'Twin Earth', we think that the *meaning* of 'good' would be the same across Earth and Twin Earth. Hence, the 'Twin Earth' thought experiment suggests that the meaning of 'good' is not determined by the supernatural environment we inhabit; we do not believe that the meaning of 'good' tracks God across all possible worlds. Therefore, the thought experiment establishes that (3) & (4) are 'open', which in turn generates a powerful challenge to Adams' position; he has to account for why semantic intuitions differ from (1) & (2) to (3) & (4). This is the central point of the chapter: *we can successfully apply the Horgan/Timmons OQA to Adams' position.*

Perhaps things will be clearer if we consider Adams' own discussion of the 'Twin Earth' experiment. Adams introduces this in developing his account of morally wrong.³⁵³ Like his account of good, he starts by considering the characteristics of morally wrong and then tries to

³⁵¹ That is, before the discovery of any differences in supernatural properties between their planets.

³⁵² I use the 'action-guiding' role - specifically, the connection with motivation, and the connection with reason - as shorthand for the 'nominal essence' of good; which I take to be the constituents of common sense morality.

³⁵³ Adams (1979).

ascertain which property, if any, has these features. He suggests that being contrary to the commands of a loving God best fits this role. Thus, he concludes that wrong is identical with being contrary to the commands of a loving God. Recognising that the 'Twin Earth' thought experiment could be a problem for his account he discusses it, attempting to show why, in fact, it does not challenge his position:

... there is still, I suppose, a possible world, w_4 , in which there would not be a loving God but there would be people to whom w_4 would seem much as the actual world seems to us, and who would use the word "wrong" much as we use it... We can even say that they would believe, as we do, that cruelty is wrong... [that is to say] that the subjective psychological state that they would express by the ascription is that same that we express.³⁵⁴

Adams claims that he is open to the possibility of Earthlings and w_4 inhabitants *meaning* the same thing by the word 'wrong'. I assume though that this is a slip on Adams' part; yet, it is surely a revealing one.³⁵⁵ It reveals Adams' desire to hold onto the intuition that if two communities share the features central to common sense morality, such as the action guiding role of moral judgements, then one would conclude that members of those communities mean the same by their moral terms. That is, even given that for one there is no God and for the other

354 Adams (1975: 117). Emphasis added. In fact, even stating these points is odd for surely the existence of God is not contingent. Perhaps then Adams ought to have argued that 'wrongness' is a *strongly rigid designator*, rather than merely a rigid designator. See Kripke (1980: 49).

355 Adams (1979: 117). Revealing in that it highlights his desire to retain the ability of Earthlings and w_4 inhabitants to be in genuine disagreement.

there is. That Adams cannot secure these intuitions becomes apparent when I clarify the implications of his account.

Given Adams' semantic account, he cannot assert, as he does, that there is *any* common cognitive content across the actual world and w_4 . To make such a claim is to miss the point of Putnam's thought experiment. The thought experiment demonstrates that if one adopts internalism about cognitive content, then Earthlings and Twin Earthlings can never have the same cognitive content; briefly, this is why. For ease, adopt a Fregean context for this discussion. First, assume that Frege is correct to claim that sense determines reference. Second, assume that cognitive content determines sense grasped (if individuals are cognitively identical and use E, then they associate the same sense with E). Putnam's argument proceeds. Suppose that two *atom-for-atom doppelgangers* Toscar (from Twin Earth) and Oscar (from Earth) visit each other's planets and both are facing a puddle. Oscar thinks (and says, pointing) 'This is water' and Toscar thinks (and says, pointing) 'This is water'. If water and twin water really are different substances, then Oscars 'water' does not have the same reference as Toscar's 'water'. And given that sense determines reference, Oscar's 'water' has a different sense from Toscar's 'water'. But we granted that Oscar and Toscar are atom-for-atom replicas; thus, what I label 'cognitive internalism' seems false since it requires the conclusion that Oscar and Toscar are cognitively identical. This is why Putnam famously claimed that 'meanings just ain't in the head!'³⁵⁶

Now apply this to what Adam's says. Oscar's *belief* that there is water [H_2O] in a glass is, as a matter of necessity, different from Toscar's belief that there is water [XYZ] in a glass. We

³⁵⁶ Putnam (1975).

can make the same point, *mutatis mutandis*, for any propositional attitude. Given that Adams thinks that 'good' is a rigid designator which tracks resemblance to God across all possible worlds we cannot hold both that his account is correct and that there is a common cognitive content between Earthlings and w_4 inhabitants. If there were no God in w_4 , then w_4 inhabitants cannot *think, hope, believe*, etc. anything about good [resemblance to God]. Hence, Adams ought to have reasoned as follows:

... there is still, I suppose, a possible world, w_4 , in which there would not be a loving God but there would be people to whom w_4 would seem much as the actual world seems to us, and who would use the word "wrong" much as we use it... We *cannot* however say that they believe, as we do, that cruelty is wrong, that is to say we *cannot* say that the subjective psychological state that they express by the ascription is the same as we express – they express different subjective psychological states.

This highlights the implausibility of Adams' semantic position. For instance, do we really want to be committed to the claim that we *cannot* say that w_4 inhabitants believe, as Earthlings do, that cruelty is wrong purely because on Earth there is a God and on w_4 there is not? It is, at least, not obvious that we do. Hence, questions (3) & (4) are 'open', and that, as such, the revised OQA successfully challenges Adams' account.

In conclusion, the OQA is not about epistemic possibility but about semantic practice – as demonstrated by our linguistic intuition. Perhaps, though, Adams could respond that the comparison with Boyd's account is too close. There is a significant difference between the

accounts such that his is more plausible. It seems to me that the only way to establish such a conclusion would be to claim that because God is beyond human knowledge a test based on closing epistemic possibility is bound to be unrepresentative. Thus, we cannot justify this version of the OQA. I find such a response unconvincing; the 'Twin Earth' thought experiment does not rely on us ever coming to know that water is H_2O . It simply supposes, as a thought experiment, that we do in fact know that water is H_2O . Hence, unless Adams can show that we can not even stipulate that Good is resemblance to God, he can not show this response to be viable. Hence, one *can* apply Horgan and Timmons' OQA to Adams' account.

However, if we had to pick between Adams' and Boyd's accounts, it is Boyd's which seems more plausible. Consider the claim that water is metaphysically identical with H_2O . Initially we can specify an *a priori* conceptual claim concerning water, which specifies water's nominal essence. The scientists then investigate, and eventually discover, *a posteriori*, the real essence which fits this specification – H_2O . Importantly, it is because the discovery is made *a posteriori* that the identity 'water is H_2O ' is *synthetic* yet necessary. Boyd's account adopts the same methodology; he starts with an *a priori* conceptual claim concerning good, and then the appropriate people investigate, and eventually discover, which homeostatic cluster of natural properties fits the role. Boyd can then make a synthetic yet necessary identity claim between a moral property and the discovered homeostatic cluster of natural properties. It seems unlikely to me that a similar analogy is available to establish Adams' account.

Just as in Boyd's account, Adams' position starts with an *a priori* conceptual claim concerning good. Once this is in place, to secure the analogy Adams would have to claim that we discover *a posteriori* the property that fulfils this role. However, talk about *a posteriori*

discovery seems out of place when talking about God. Surely, we do not think that an empirical investigation of the world could lead to the *a posteriori* discovery that good is resemblance to God. More likely, in such a case we would talk about *revelation* and *insight* rather than discovery. Obviously, the epistemological status of beliefs about revealed truths concerning God is controversial. However, it is clear that such claims are closely connected to *faith*, and that as such we do not have *a posteriori* knowledge of them. I suggest that if we expose two people with the same physiological capabilities to the same descriptive information then one may and one may not 'discover' that good is to resemblance to God. We are not forced to conclude that in this instance one must be mistaken, or needs to 'properly' consider the descriptive information. It is more correct to claim that people know good is resemblance to Good, *a priori*. However, *a priori* conceptual claims about 'good' make up its 'nominal essence'. Hence, good being resemblance to God is something that Adams ought to think figures in the characteristics that 'fix the reference' of good, rather than being something that is discovered *a posteriori*. There then seems to be a fundamental problem for Adams' semantic framework, even before we consider the OQA.

§3. Conclusion

Philosophers have often thought that the 'openness' of open questions reflect human's cognitive limitations. That is, we find questions open because we do not, or cannot, fully grasp the nature of moral properties. Hence, Iris Murdoch claims:

Good is indefinable not for the reasons offered by Moore's successors, but because of the *infinite difficulty of the task of apprehending a magnetic but inexhaustible reality*.³⁵⁷

Similarly, Frank Jackson claims that he can reject the OQA because it relies on moral properties being 'mysterious' and 'unknowable'; where in fact they are not.³⁵⁸ However, it is our semantic practice which grounds the openness of open questions and hence it is a mistake to think the OQA is based on epistemic possibility.³⁵⁹ Consequently, Adams' account is not at an advantage over naturalism by claiming that good is a supernatural property. In fact, no account can adequately respond to the OQA by arguing that moral properties cannot be known.

357 Murdoch (1970: 42).

358 Jackson (1998: 151, 167).

359 See chapter IX

(VIII)
DISPOSITIONAL ACCOUNTS OF MEANING AND THE OQA

The aim of this chapter is to ascertain whether the OQA challenges a dispositional account of meaning, I argue it does. The chapter does not contribute to the general enquiry into moral naturalism but addresses a side issue.

§1 Sophisticated Dispositionalism

A dispositional account of meaning can be 'simple' or 'sophisticated'. The latter is the stronger position and will be the focus of this chapter. However, to better understand it I briefly outline simple dispositionalism.

Dispositionalism claims that there is a natural fact – a categorical base - in virtue of which ascriptions of meaning, such as "Andy means addition by '+'", are either true or false.³⁶⁰ It argues then that meta-semantic claims are truth-apt and have truth-makers. Simple dispositionalism claims that such a sentence is true when Andy is disposed to respond to the relevant arithmetical queries with the sum of the numbers involved. For example, "Andy means addition by '+'" is true when Andy answers '4' when asked to compute '2+2', answers '10' when asked to compute '5+5', etc. The problem with such a simple account is that it cannot distinguish between competence and performance.

³⁶⁰ Interestingly, Stuart (1998) argues that dispositions need not be thought of as natural facts. However by 'disposition' I mean a categorical base which brings about certain behaviour.

Andy is not omniscient, and has not always given the correct answer in summing numbers. So, mistakes in calculation were part of his performance in summing numbers. Thus, according to the simple dispositionalist "Andy means addition by '+'" is true in virtue of Andy not only answering '10' to '5+5', but also answering '8' to '2+1' and '100' to '56 + 43', etc. Of course, intuitively we want to exclude these two answers but this is exactly what simple dispositionalism cannot do - it fails, as it cannot allow for mistakes. Hence, I reject simple dispositionalism.

If there were a way of distinguishing between competence and performance then such a simple refutation of dispositionalism would not be available. By introducing an ideal condition clause, the sophisticated dispositionalist claims that this is possible. It is not that "Andy means addition by '+'" is true when Andy is disposed to respond to the relevant arithmetical queries with the sum of the numbers involved. It is true when he is so disposed to do so under ideal conditions, the basic thought being that Andy would not make mistakes; he wouldn't answer '8' when asked '2+1', under ideal conditions.³⁶¹ We can outline sophisticated dispositionalism as follows:

(SD) Andy means addition by '+' = Given ideal conditions (of the environment and of Andy), when asked to compute 'x+y', Andy is disposed to produce the sum of the two numbers.

If (SD) is correct then there is a natural fact in virtue of which "Andy means addition by '+'" is either true or false, and the distinction between competence and performance can be preserved.

³⁶¹ This is controversial, but I assume it is true.

Saul Kripke has suggested a number of reasons why sophisticated dispositionalism fails.³⁶² However, I won't discuss all these. I shall focus on the 'normativity argument' and argue that (i) if it were successful then it would challenge sophisticated dispositionalism; but (ii) its crucial premise relies on a traditional account of the OQA and hence fails. In the next section, I argue there are two other forms of the OQA that justify the normativity argument.

Suppose that Andy means addition by '+'. What is the relation of this to the question: how will Andy respond to the problem '68+57'? The dispositionalist gives a descriptive answer: if Andy means addition by '+', then Andy *would* answer '125'. However, Kripke argues that the relation of meaning to future action is normative not descriptive.³⁶³ If Andy intends to accord with his past meaning of '+', he *should* answer '125'. The main ideas here can be set out:

(1) The best the dispositionalist can say is this: under ideal conditions Andy *would* respond with such and such a number to each addition sum.

(2) However, Andy's meaning addition by '+' determines what answer he *should* give, not what answer he would give. Meaning facts are normative facts, while dispositional facts are descriptive facts.³⁶⁴

(3) Descriptive facts cannot license evaluative claims.³⁶⁵

362 See Kripke (1980).

363 Kripke (1982: 24 – 37).

364 By 'meaning fact' I mean the facts that determine which objects satisfy a predicate, as meant by a speaker at a time. Furthermore, the satisfaction conditions of a predicate are determined by a property to which the predicates bear a certain relation. Whether an object satisfies a predicate is determined by whether the object exemplifies the property to which the predicate bears this relation. Zalabardo (1997: 468).

(4) It is not the case that Andy's meaning addition (or some other function) by

'+' is constituted by his ideal dispositions with the '+' sign. [From (1), (2) & (3)]³⁶⁶

This argument is valid; however, does this mean that sophisticated dispositionalism ought to be rejected? That depends on whether the argument is sound.

For the sake of argument I assume that (1) is true. (2) claims that dispositions are descriptive facts, this also seems correct. After all, we have already granted that dispositions are natural facts.³⁶⁷ However, (2) also claims that meaning facts are normative facts, or meaning facts determine what answer should be given. This too seems correct; after grasping the meaning of a term we think that this somehow 'tells' us how we ought to apply the term in the future. What does the 'ought' mean in this instance? It seems to be, at least, something like this: if you want to apply the term correctly, then use it in *this* way. The meaning of a predicate 'contains' the conditions of correct use. As Boghossian puts it:

Suppose the expression "green" means *green*. It follows immediately that the expression "green" applies *correctly* to *these* things (the green ones) and not to *those* (the non-greens). The fact that the expression means something implies, that is, a whole set of *normative* truths about my behaviour with that expression: namely, that my use of it is correct in application to certain objects and not in application to others... The normativity

365 One cannot justify inferring evaluative conclusions from descriptive premises.

366 Note: I will use 'ideal disposition' as shorthand for 'disposition under ideal conditions'.

367 The categorical bases which when appropriate conditions obtain manifest behaviour.

of meaning turns out to be, in other words, simply a new name for the familiar fact that ... meaningful expressions possess conditions of *correct use*.³⁶⁸

At this stage I do not want to engage in the discussion concerning whether this is the most useful way of understanding normativity, but merely assume that (1) and (2) are true. The soundness of the normativity argument then depends on (3). Why might one think that one cannot justify drawing evaluative conclusions from descriptive premises?³⁶⁹

One answer may lie in the fact that dispositions are finite, whereas the question of how we would use an expression, in principle, should receive an answer in infinitely many cases: conditions of correct use cover a potentially infinite number of cases.³⁷⁰ However, it is not clear why this fact would justify premise (3). This is demonstrated by making explicit the implicit premises in this way of proceeding:

... that the finiteness of dispositions follows from the descriptive character of dispositional facts, and... that it is in virtue of its normative character that the question of how I should apply a predicate has to receive an answer in infinitely many cases.³⁷¹

It is unclear whether either of these claims is true.³⁷² Hence, I assume that the finiteness of dispositions cannot justify (3).³⁷³ Jerry Fodor suggests that the most likely account of why (3)

368 Boghossian (1989: 129).

369 In terms of dispositionalism, why think that 'the descriptive character of claims about speaker's dispositions and the normative character of claims about how they should use linguistic expressions invalidates inferences from the former to the latter'? Zalabardo (1997: 471).

370 The ideas concerning finiteness are taken from Zalabardo (1997).

371 Zalabardo (1997: 472).

is thought to be true is the OQA.³⁷⁴ He argues that a form of an ‘open question’ argument supports the reasoning behind such claims, and it is precisely for this reason that the normativity argument fails to challenge his version of semantic naturalism.³⁷⁵ He writes:

In short, I’m not clear how – or whether – “open question” arguments can get a grip in the present case. I am darkly suspicious that the Kripkensteinian worry about the normative force of meaning is either a non-issue or just the reduction issue over again; anyhow, that it’s not a *new* issue.³⁷⁶

If Fodor is correct, then what supports the claim that we cannot draw evaluative conclusions from descriptive premises is the OQA. How a speaker would apply a predicate under ideal conditions cannot be the fact that constitutes how he should apply it, because we can meaningfully ask whether the applications that he would endorse under ideal conditions are the ones that he should endorse. However, Fodor suggests that if this is what supports the normativity argument then it fails, because:

The problem is that proponents of naturalistic reductions of semantic notions *see their task as on a par with other theoretic reductions*, such as the identification of water with

372 As Zalabardo cites, Fodor (1990: 94 – 95), argues that dispositions need not be thought of as finite.

373 Equally, the fact that sophisticated dispositionalism has problems specifying ideal conditions non-circularly will not show why we ought to accept (3). The standard normativity argument is meant to challenge sophisticated dispositionalism even if a specification of ideal conditions were available.

374 Fodor (1990). Zalabardo (1997) argues that, in the end, (3) is nothing more than Hume’s claim that one cannot derive and ‘ought’ from an ‘is’. Hume (1778: 469). However, as noted in chapter (II) this understanding of the OQA would not be in line with Moore’s OQA.

375 Fodor (1990: 135).

376 Fodor (1990: 136)

H₂O or heat with kinetic energy... No doubt we can meaningfully ask whether water is H₂O, but the meaningfulness of this question doesn't undermine the identification of water with H₂O – the claim that whether something is water is determined by whether it has that molecular structure. Similarly, semantic naturalists would concede that we can meaningfully ask whether a speaker should ascribe a predicate to the objects to which he would ascribe it under certain conditions. But the meaningfulness of this question, they would contend, doesn't undermine the claim that how the speaker should apply the predicate is determined by how he would apply it under those conditions.³⁷⁷

The sophisticated dispositionalist sees his account 'on a par with theoretical reductions': Andy means addition by '+' is the fact that when asked to compute 'x+y', Andy is disposed, under ideal conditions, to produce the sum of the two numbers. But that 'Andy means addition by "+"' does not mean, 'when asked to compute "x+y", Andy is disposed, under ideal conditions, to produce the sum of the two numbers'. The OQA fails to support the essential premise in the normativity argument as it assumes that property identity requires synonymy. It therefore fails to challenge sophisticated dispositionalism.

In the next section I argue that there are two versions of the OQA which can be applied to non-analytic naturalism, and hence that they support the premise in the normativity argument. I claim that as such we ought to accept the conclusion that it is not the case that Andy's meaning addition (or some other function) by '+' is constituted by his ideal dispositions with the '+' sign.

³⁷⁷ Zalabardo (1997: 476 – 7). I argue against this way of understanding the task of explaining predicate reference in the next chapter.

§2 Can the OQA Support The Normativity Argument?

I argue that there is justification for (3) via the OQA, and hence the normativity argument is both sound and valid. To understand this we need to be aware that what the OQA shows is that dispositional facts cannot *justify* responses; it is the connection between normativity and justification that renders the OQA a genuine challenge to sophisticated dispositionalism.

For the sake of argument, assume that for a meaning fact to justify my response to a addition sum, it must *tell me* how to respond. The argument I present is that the OQA suggests (3) is true because our dispositions cannot tell us what applications to endorse. To see why this is the case consider another question: 'why would typing '2+2' in a calculator be a justified procedure for answering the sum, whereas rolling a die would not be?' Whether a given procedure for answering additions questions is justified can be expected to depend on which facts determine how addition questions should be answered, and on how the procedure is related to these facts. There are two ways of understanding this relation.

First, if we start from an account concerning which facts determine how we should answer '2+2', then from this we can generate a criterion of adequacy for the procedures that we may want to use for answering 'what is 2+2?' one that, for instance, would rule out die rolling and 'rule in' calculator use.

Second, if we start from a hypothesis concerning which procedures are justified, then this will generate a criterion of adequacy for accounts of which facts determine how '2+2' should be answered.

... the facts that play this role will have to be related to those procedures in such a way as to render them justified.³⁷⁸

Therefore, on this second way of proceeding, how the facts that determine how to answer '2+2' are related to using a calculator, for instance, will help us answer why this procedure is justified, whereas die rolling is not.

For the sake of argument I assume that the correct approach in discussing justification in relation to sophisticated dispositionalism is the latter kind of criterion. In order to assess sophisticated dispositionalism we need to start from a hypothesis concerning which procedures are justified, and this will generate a criterion of adequacy for accounts of which facts determine how '2+2' should be answered. That is, the fact that determines how Andy should answer '2+2' has to be related to the procedures that Andy actually employs for answering these questions in such a way as to render the procedure justified.

One plausible account argues that a procedure for answering '2+2' is only justified if it involved *conscious engagement* with the fact that determines how addition questions should be answered. The thought behind this is an internalist account of epistemic justification, a view summed up by John Greco:

... the conditions for justification must be appropriately internal to the knower's perspective. Roughly, something is internal to S's perspective so long as S is aware of it or could be aware of it merely by reflecting.³⁷⁹

378 Zalabardo (1997: 479).

Accepting internalism, we can set up an OQA that challenges sophisticated dispositionalism. The sophisticated dispositionalist claims that the fact that determines how questions such as '2+2' should be answered by Andy is the disposition he would have under ideal conditions. Hence, Andy's response to '2+2' is justified if it involved conscious engagement with ideal dispositions. However, it is an open question whether the way that Andy responds to addition sums is the way he would under ideal conditions. Hence, the dispositionalist is forced to conclude that facts about ideal dispositions cannot determine how Andy should respond to '2+2'. This version of the OQA is as follows:

(1*) The procedures that Andy uses for deciding how to respond to '2+2' must be justified.

(2*) The procedures that Andy uses for deciding how to respond to '2+2' would only be justified if they involved conscious engagement with the facts that determine which response Andy should give.

(3*) If Andy decided how to respond to '2+2' by considering what Andy would do under ideal conditions, then it would be a closed question for Andy whether under ideal conditions he would in fact respond in the way he did.

(4*) It is an open question whether under ideal conditions Andy would in fact respond in the way he did.

(5*) Andy decides how to respond to '2+2' without considering whether he would do so under ideal conditions (from (3) and (4))

(6*) If facts about how Andy would respond under ideal conditions determine which response he should give, the procedures that Andy use for deciding how to respond would not be justified (from (5) and (2))

(7*) Facts about how Andy would respond to '2+2' under ideal conditions do not determine how I ought to respond.

Thus, it seems that Fodor was wrong to dismiss the OQA so quickly. Even given that the sophisticated dispositionist bases his account on identities such as ' H_2O is water', there is a challenge. Hence, using the OQA, we can assert that the following argument is sound and valid:

(1) The best the dispositionalist can say is this: under ideal conditions Andy *would* respond with such and such a number to each addition sum.

(2) However, Andy's meaning addition by '+' determines what answer he *should* give, not what answer he would give. Meaning facts are normative facts, while dispositional facts are descriptive facts.

(3) Descriptive facts cannot license evaluative claims.

(4) It is not the case that Andy's meaning addition (or some other function) by '+' is constituted by his ideal dispositions with the '+' sign. [From (1), (2) & (3)]

There are two main ways that the sophisticated-dispositionalist could respond to this. First, he could reject (1*) and claim that the procedure that Andy uses to decide how to answer '2+2' need not be justified. This however seems very radical; it holds that all responses are 'stabs in the dark', and that we might as well pick a number from a hat as type the sum in the calculator. If the dispositionalist is forced to accept this then, I suggest, this is a good reason to reject his account.

More promising would be for the dispositionalist to claim that Andy does not require conscious engagement with the facts that determine what answer he should give for his response to be justified. Given this, it is – the argument would continue – no surprise that it is an open question to Andy whether the way that he responded was the way he would have done if he were under ideal conditions. After all, it is *not transparent* to him that his procedures track the disposition. To adequately assess if we can reject internalism requires an in depth discussion of

epistemic externalism and related issues – something I do not enter into. However, I think there are reasons – highlighted by the OQA – why it is problematic to reject (2*) and to claim that Andy's responses are justified because his procedures for responding to addition sums *reliably track* the facts that determine how the questions ought to be answered.

The version of the OQA which will support this claim challenges the assumption that the dispositionalist can, without complication, argue his position is 'on a par with theoretical identities.' This version relies on my discussion of Horgan & Timmons in chapter (VII). Recall that they argued that we could generate an open/closed distinction when considering theoretical identities. Horgan and Timmons suggest that linguistically competent agents who reflect on 'Twin Earth' thought experiments would claim that these questions are 'closed':

- (5) *Given* that the use of 'water' by humans is causally regulated by the natural kind H_2O ,
is liquid L, which is H_2O , water?
- (6) *Given* that the use of 'water' by humans is causally regulated by the natural kind H_2O ,
is liquid L, which is water, H_2O ?³⁸⁰

If the analogy between sophisticated dispositionalism and theoretical identities holds, then we should be able to write suitably altered question that are also closed. Here is my attempt at this.

380 This is as opposed to the questions (i) Liquid L is H_2O , but is it water? And (ii) Liquid L is water, but is it H_2O ?, which are 'open'.

(7) *Given* that the use of “Andy means addition by ‘+’” by humans is causally regulated by Andy being disposed to produce, when asked to compute ‘ $x+y$ ’, the sum of the two numbers under ideal conditions. Is Andy’s disposition identical with Andy’s meaning addition by ‘+’?

(8) *Given* that the use of “Andy means addition by ‘+’” by humans is causally regulated by Andy being disposed, when asked to compute ‘ $x+y$ ’, to give the sum of the two numbers; is Andy’s meaning addition by ‘+’ identical with Andy’s disposition?

If sophisticated dispositionalism is correct then a competent speaker who reflects on (7) & (8), and who taps into his own linguistic competence ought to judge that the answer to these question is obviously ‘yes’. I suggest, however, that they would not, and that a suitably altered version of the Twin Earth demonstrates that our semantic intuitions do not ‘close’ (7) & (18). This shows that the dispositionalist cannot, without qualification, claim his position is ‘on a par with theoretical identities’. Consider a ‘Semantic Twin Earth’ thought experiment.

On Earth Andy exhibits behaviour that common sense takes to manifest *meaning addition* by “+”, and on Twin Earth Twin-Andy exhibits exactly the same behaviour. However, they differ in their disposition; that is, in the ‘categorical base’ underlying their identical behaviour. Call the base on Earth *E*, and on Twin Earth *TE*. Now, assume that an Earthian visits Twin Earth and on observing Andy’s behaviour say ‘Andy means addition by ‘+’’. There are two possible ways of describing this situation.

If sophisticated dispositionalists are correct we would expect our intuitions to be as follows. When the Earthian makes his pronouncement on what Andy means, he makes a mistake, for he refers to the categorical base *E* and not *TE*. Moreover, if the Earthian talks with the Twin Earthian about whether Andy means addition by '+', they are talking at 'cross-purposes'. They mean different things by their claim that 'Andy means addition by '+'.

The second way of describing things is that given the behaviour of Andy and Twin Andy is the same, they both exhibit behaviour that common sense takes to manifest *meaning addition by '+'*. Then the Earthian and Twin Earthian mean the same by 'Andy means addition by '+'', and there can be genuine disagreement between Earthians and Twin Earthians.

I judge that the second way is the most natural way of describing the thought experiment, as 'Andy means addition by plus' is warranted so long as the behaviour is found.³⁸¹ Hence what this suggests is that we are not predicating any particular categorical property in predicating it; that is, even *given* that our '+' talk is causally regulated by some underlying physical property *E*, it is still an open question whether the property *E* = the property of meaning addition by '+'. If this is the case then it 'gets things wrong' for the semantic naturalist who sees his task on a par with other theoretical reductions. If they are 'on a par' then reflection on this thought experiment ought to generate intuitions analogous to those generated in Putnam's thought experiment. That is, that Semantic Twin Earthians and the Earthians mean something different.³⁸² It seems then

381 Even so, the meaning of 'Andy means addition by '+' cannot be reduced to Andy's behaviour. See the next chapter for a rejection of such a reductionist strategy.

382 There is another puzzling side to this, one that I do not discuss. If 'Andy means addition by '+' can be analysed in terms of Andy's possessing some intentions and beliefs about the usage of '+', then we have an identity between meaning that and having these intentions/beliefs. Yet, I doubt that the OQA challenges this identity: that is, between meaning something and having such-and-such beliefs and intentions. Meaning something is indeed having certain beliefs and intentions. So does the proposition that X has these intentions and beliefs entail some normative proposition? That is, if 'meaning is normative' then

that the meaning of 'Andy means addition by '+' depends on the similarities in the behaviour across worlds, rather than identity in the categorical bases which are manifest in such behaviour.³⁸³

Thus, one can state a revised OQA that challenges the sophisticated dispositional account:

ROQA: *Revised Open Question Argument*: Questions of the form (9) and (10) are open questions.

That is, closing the epistemic possibility does not 'close' (9) and (10), they remain 'semantically open'. If this is correct, and the sophisticated dispositionalist cannot respond to it, then they have to accept the truth of (3). Furthermore, given that the other premises are true, and the argument is valid, they would have to accept the conclusion of the normativity argument.

(1*) The best the dispositionalist can say is this: Under ideal conditions I would respond with such and such a number to each addition sum.

(2*) However, my meaning addition by '+' determines what answer I should

are certain beliefs and intentions normative? It seems plausible to answer 'yes'. In which case the OQA, if sound, would challenge the identity between having those beliefs and intention and being in some physically characterised state.

383 But this does not mean that one can reduce 'meaning' to such similarities.

give (i.e. the answer in accord with the function I mean), not what answer I would give. In other words, meaning facts are normative facts, while dispositional facts are descriptive facts.

(3*) From the revised OQA, a set of descriptive facts cannot entail that another set of facts (here, my uses of the '+' sign) are correct or incorrect. In other words, descriptive facts cannot licence normative facts.

Therefore,

(4*) It is not the case that my meaning addition (or some other function) by '+' is constituted by my (ideal) dispositions with the '+' sign. [From (1*), (2*) & (3*)].

Notice also that Horgan & Timmons' adapted argument does not just hold for dispositional accounts. Specifically, OQA would challenge any semantic naturalist position according to which semantic terms bear some relation *R* to certain natural properties, and semantic terms refer to the natural properties to which they bear this relation *R*. In such cases, one could reconstruct a similar 'Semantic Twin Earth' thought experiment by rewriting the causal role in terms of the relation *R*. For example, Jerry Fodor's asymmetric dependency

account of meaning would stand challenged by this approach.³⁸⁴ As with the discussion in the moral case there are three possible responses that the semantic naturalist could make to this version of Horgan & Timmons's OQA.

First, they could deny that Horgan & Timmons's account describes a genuine possibility, and so hold that our intuitions have no evidential weight. This response fails, as it is too strong. It would require defending the conceptual impossibility of such thought experiments. Given the 'Semantic Twin Earth' scenario is intelligible, this would incur an overwhelming burden of proof.

Second, they can argue that although (9) & (10) are 'open', there is no guarantee that they will stay this way; a categorical base may end up 'closing' them. Apart from an unrealistic optimism, this response fails because it disregards the fact that Horgan & Timmons's OQA is a general argumentative *strategy*. The naturalist cannot argue against them by simply asserting – without further argument – that eventually their position will be vindicated.

Third, they can argue that people's intuitions about 'Semantic Twin Earth' are simply mistaken, and hence do not undermine causal semantic naturalism. After all, brute intuitions do not settle the matter by themselves. However, responding like this would again place a great burden of proof on the semantic naturalist. He would have to demonstrate why the intuitions concerning 'Semantic Twin Earth' could be categorised differently from those concerning 'Twin Earth'. He would have to account for why people's meaning-intuitions about meaning are so strong and widespread, even though they are allegedly mistaken. I suggest none of these three responses would be satisfactory.

384 Fodor (1990)

§3. Conclusion.

I argued that the most reasonable dispositionalist accounts is a sophisticated one, as this avoids problems to do with competence and performance. However, this position is threatened by the normativity argument, though in its original form, this seems unable to challenge dispositionalist accounts that see their identities 'on a par with theoretical identities'. However, even granting this, two forms of the OQA do present a challenge. This in turn means that the sophisticated semantic dispositionalist cannot dismiss the normativity argument but is challenged to accept its conclusion: it is not the case that my meaning addition (or some other function) by '+' is constituted by my (ideal) dispositions with the '+' sign. I conclude that when applied to dispositional accounts of meaning, the OQA is a genuine argumentative device.

(IX)

CONCLUSION AND SUGGESTION

In §1 of this chapter I draw the main conclusions of the previous eight chapters. In §2 and §3 I briefly give an account of the moral theory I think we should adopt in light of these. In §4 I draw some conclusions.

§1. Conclusion.

I started the thesis with a discussion of G. E. Moore's OQA.³⁸⁵ It quickly became apparent that Moore's presentation of his ideas are confused. His open question argument relies on his conception of analysis, phenomenology, metaphysics and epistemology. All of which were problematic. I then used three modern interpretations of the OQA to show that although Moore's OQA should be abandoned, there is something of value in his argument; in the words of Darwall, Gibbard and Railton, I argued that Moore was 'on to something'.³⁸⁶ The main outcomes of these three accounts are first that the 'open question' argument is, at best, a challenge and not a 'knock-down' argument. We should note that people we judge to be normal, do, in fact, find the argument compelling. Second, that the OQA only challenges those accounts which hold that property identity requires synonymy. Out of the three accounts discussed, I developed Darwall, Gibbard and Railton's (DGR).³⁸⁷ Their proposal relied on there being a conceptual connection between moral judgement and motivation, and a conceptual connection

³⁸⁵ Moore (1993).

³⁸⁶ Darwall, Gibbard and Railton (1992).

³⁸⁷ Darwall, Gibbard and Railton (1992).

between something being right and practical reason. I argued that their former argument fails, as there is no such conceptual connection between judgement and motivation. However, considering the link between them did pose a challenge to the analytic naturalist. That is, for the analytic naturalist to secure the truth of his reduction he would have to demonstrate that naturalistic judgements have the capacity to motivate in and of themselves. I suggested that it is hard for the naturalist to do this, and that this is the first account of why the OQA is a genuine argumentative device. I then moved to discuss DGR's second claim. I found this argument more convincing than the last, because there is a conceptual link between something being right and there being reason to do it. This conceptual connection thus has to be respected by the analytic naturalist. However, our intuitions suggest that there is no conceptual connection between judging that ϕ -ing would produce N and the judgement that one has any practical reason to ϕ . Thus, the second challenge from the OQA is this. Either the analytic naturalist has to demonstrate why there is no conceptual link between something being right and practical reason, or he has to account for why people have the convictions they do.

I then turned my attention to non-analytic naturalism, which is thought to be immune to open question arguments. Non-analytic naturalists argue that the open question argument relies on the premise that property identity requires synonymy, but that their account shows this to be false. I considered two recent non-analytic positions, Peter Railton's and Richard Boyd's.³⁸⁸ I argued that there is a version of the OQA which challenges both these accounts. The OQA that challenges Railton's account turns on the fact that his naturalisation of normativity is incomplete. Questions concerning Railton's account are 'open' because his account fails to

388 Railton (1986a), (1986b) and (1989), and Boyd (1988).

recognise the role of ideals for people. Possible selves inform our decisions, and the notion of possible selves requires the ability to change personality. This is something that Railton's account rejects.

I rejected Boyd's account because once we close epistemic possibility, the semantic intuitions which are meant to support Boyd's account fail to do so.³⁸⁹ Furthermore, if we consider our referential intuitions with moral terms this demonstrates that identities between moral properties and natural properties are not like theoretical identities. The third and fourth conclusions were then that the OQA challenges methodological, substantive naturalism, and it challenges those naturalistic accounts which draw an analogy with theoretical identities.

I then considered supernaturalism. The thought was that perhaps such an account could avoid these conclusions. Any position which claims that moral properties are beyond our knowledge will leave questions concerning the account open. However, by closing epistemic possibility and considering referential intentions I demonstrated that this position is also challenged by the OQA. This led me to the fifth conclusion: the OQA is not to do with epistemological possibility, but to do with our semantic practice.

Then I discussed another naturalist account, but this time in terms of semantics. I was interested in whether the normativity of meaning meant that the OQA challenges sophisticated dispositionalist accounts of meaning. I concluded that it does. The main reason is that there is no way of justifying the sophisticated dispositionalist claim that their identities between semantic facts and dispositional facts are on a par with theoretical identities. This, then, was my sixth conclusion.

³⁸⁹ The same argument also applies to Railton's account.

In light of these six conclusions, arrived at via a discussion of the OQA, there are two options. Perhaps we ought to develop a more sophisticated naturalist account, or perhaps we should adopt a non-reductive moral realism, or a non-cognitive account. Maybe we will be able to revise and develop such positions so as to respect all that has been said thus far. I do not discuss this possibility here. The other way of proceeding is to try to give an account that avoids reductionism and realism, but respects the six conclusions noted. I consider this option and suggest that the lesson of the OQA is that we ought to adopt moral nominalism, or so I shall argue.

§2. Nominalism

This section is a tentative suggestion as to how we might develop a position in response to the conclusions drawn thus far. I call such a position ethical nominalism. First, I discuss an account of predicate applicability that supports ethical nominalism and second, I argue that such a position should explain our application of moral predicates in terms of those spontaneous inclinations to apply them that arise from our conscience.

There is a version of ethical nominalism which is clearly false. It holds that moral predicates are applicable to states of affairs just in case speakers feel inclined to apply them. Its implausibility lies in the fact that such a view fails to take into account two fundamental aspects of our moral practice - features that the OQA makes perspicuous. Specifically, a predicate as meant by a speaker can be applicable to something even though the speaker is not inclined to apply that predicate to that thing; and a predicate may not be applicable to something even though a speaker may feel inclined to apply it to that thing. Henceforth, I eschew this version of

'nominalism' and attempt to develop another. The view I discuss can be summed up as follows: the applicability of moral predicates can be *explained* solely in terms of speakers' feeling inclined to apply predicates to some situations and not to others, but we cannot *define* moral predicates in terms of these inclinations. I argue that such a position can demonstrate why moral predicates seem beyond definition - as highlighted by the OQA - and that because such an account does not rely on moral properties it avoids epistemological and metaphysical worries associated with moral realism. First, though I develop an account of predicates applicability which supports ethical nominalism.

G. E. Moore was right, but for the wrong reasons, when he wrote that 'good' is indefinable.³⁹⁰ Discussion concerning why 'good' cannot be defined has been guided by an assumption: to resist moral realism one has to show that moral predicates can be defined in terms of speaker's inclinations to apply them: for example, in terms of their inclinations to apply 'good' to that which we desire to desire. However, I have argued that a proper understanding of the OQA undermines the idea that moral predicates can be reduced in such a way. It is no surprise, then, that moral realist positions have emerged, and that, for instance, such positions compare moral terms with natural kind terms. After all, philosophers such as Saul Kripke and Hilary Putnam have provided good reason to conclude that we cannot define natural kind predicates, and this is supposedly, what we are looking for in our moral predicates.³⁹¹ Whether 'water' is applicable to an object is not determined by what verdict on the matter our 'water' applying procedures support. Its applicability is not dependent on whether 'liquid', 'clear',

³⁹⁰ Moore (1903).

³⁹¹ Kripke (1980) & Putnam (1975).

'tasteless', etc. are applicable to it. We cannot give *a priori* necessary and sufficient conditions for natural kind predicates. Hence, given the assumption - accept *either* reductionism *or* realism - philosophers give a realist account of the applicability of natural kind predicates. Whether 'water', 'salt', etc are applicable depends on the presence of certain properties, H₂O, NaCl etc. Hence, given that: (i) the OQA shows that 'good' is not definable; (ii) our practice with moral predicates is similar to our practice with natural kind predicates; (iii) we hold the assumption that we accept *either* reductionism *or* realism, then we should adopt a form of moral realism. The OQA forces us to account for the applicability of 'good' by postulating a property whose presence in, or absence from, states of affairs determines its applicability.

What this clarifies is that we can resist realism whilst respecting the OQA by rejecting the assumption that we must accept *either* reductionism *or* realism so construed. Arguably, we could accept the claim that 'good' is indefinable whilst avoiding the metaphysical and epistemological worries surrounding moral realism.

Consider one attempt to define 'good'.³⁹² One way to ascertain whether something is good is to ask whether it would give us pleasure. It is clear that not everything that gives us pleasure is good; hence, one would introduce other factors to make the definition more plausible. In that way, we may arrive at the view that the 'good' is whatever gives us pleasure when we are under certain conditions. However, I argued above that this commits us to an unacceptably strong conclusion. It is surely not a *contradiction* to apply 'good' to situations where these conditions do not apply. For instance, say we define 'good' as 'that which gives pleasure to those who are statistically psychologically normal'. Do we really want to conclude

³⁹² The position in the example is naïve, but the point holds.

that there is only one answer to the question 'does the fact that the majority of people gain pleasure from something make that thing good?' that avoids *contradiction*? Surely, we do not. Perhaps, then, we would attempt to introduce some other conditions into the definition: for example, idealised epistemology,³⁹³ defining 'good' as 'that which gives epistemologically ideal agents pleasure'. This seems more plausible than the last account but, ultimately, it is not satisfactory as it is still an open question whether something which gives some epistemologically ideal person pleasure is in fact good, or so I have argued.³⁹⁴

Of course, there are other ways the reductionist can, and does, respond to this type of worry. However, the repeated application of the OQA to more complex definitions highlights a problem for reductionism. The reductionist attempts to define 'good' as meant by an agent, in terms of other predicates of his which are connected to 'good' by further beliefs of his. However, what the OQA demonstrates is that a definition is only successful when a *belief* connecting 'good' with that definition has the feature that it cannot be given up without thereby changing the applicability conditions of 'good'; call such beliefs *foundational*. For example, we cannot give up the belief connecting 'good' with 'what an epistemologically ideal agent would gain pleasure from', without thereby changing what counts as the applicability conditions for 'good'. My suggestion is that what the OQA highlights is that there can be *no* foundational beliefs connected with 'good'. There are no beliefs that connect 'good' with a definition, which when given up would change the application conditions of 'good'.

393 This is the way that Railton proceeds. See chapter V.

394 See specifically chapter VII. Here I argue that the open question challenge does not rely on our epistemic status.

Notice also that the problem is not that a particular agent has none of these foundational beliefs in relation to 'good'. *Any* beliefs are going to be equally revisable. Furthermore, notice that even if we are positive we have such foundational beliefs this is not enough.³⁹⁵ Defining 'good' in terms of the predicates to which it is linked by these beliefs may not result in a satisfactory account of the applicability of the predicate because our beliefs about which beliefs count as foundational are themselves not foundational. Assume that we arrive at a number of beliefs we thought to be foundational beliefs involving good; it is always possible that we will decide that they are in fact not foundational, and that we should give them up.

Assume, then, that the OQA demonstrates that there are no foundational beliefs connected with 'good'. As such, the realist may conclude that his position is vindicated. A property that determines the applicability condition of 'good' need not coincide with any of the properties that determine the applicability condition of other predicates in any proposed definition. 'Good' might, for instance, be a rigid designator referring to such a property irrespective of which beliefs we hold. This, though, does not avoid the challenge of the OQA. Specifically, it fails to take in to account the difference in referential intentions between moral predicates and natural kind terms.³⁹⁶ In light of these points, I discuss whether another account could capture the fact that 'good' cannot be defined.³⁹⁷

The OQA shows that our practice with 'good' is dynamic. At any one point in time, we will have a criterion for answering the question 'is this good?' Moreover, we do not arrive at

395 For instance, Jackson is positive that he has such beliefs. See my discussion of Jackson in chapter IV.

396 See specifically chapter VI and my discussion of Richard Boyd's Moral Realism.

397 It is unclear to me whether we should characterise this position irrealist or not. The issues surrounding the characterisation of realism are highly complex spawning a vast amount of literature. Simply put, though, a moral 'irrealist' rejects a metaphysically robust account of moral properties. The paradigm example of the type of metaphysical outlook the moral irrealist rejects is Robert Adams' account discussed in chapter VII. Perhaps a useful way of putting my position then is not irrealist, but anti-Platonist.

such a criterion for no reason; we are informed by other beliefs about what counts as a reliable indicator of something being good. In light of these beliefs, we may change our criterion and thus how we attempt to answer the question 'is this good?' Yet, we do not think our new criterion changes the *meaning* of the question; we believe the new answers to the questions we arrived at with our new procedure are better verdicts on the *same* subject matter. Say, for example, that the criterion we used for answering the question 'is this good' was to ask our self 'Would Jesus do this?' We have adopted this criterion because we think it is a good indicator allowing us, we believe, to dwell on God's will. However, say that we loose our faith and come to believe that the best criterion to decide on what counts as good is whatever gives us most pleasure. Hence, we decide that the best criterion is not asking 'what would Jesus do', but asking 'what would give us most pleasure?' As such, we have adopted a new criterion for deciding what is good, based on our new belief about what counts as a good criterion. Moreover, we do not think that, in light of our hedonistic criterion, we are now answering a different question; we think that the meaning of the question 'is this good?' remains the same. The point is that any belief concerning which criterion is most appropriate is revisable.

However, we do at least think our verdicts on 'is this good?' can improve. From our standpoint, within our practice with 'good', we think that we can improve at detecting what is good. However, just because we have made revisions it does not mean they are revision in the 'right direction'.³⁹⁸ After all, later we may conclude that we are actually worse now at detecting good than we were at some time in the past. In fact, this outcome is as possible as our concluding that a new way of deciding whether something is good is an improvement. What is

398 See Quine (1951).

more, the same goes for our belief that a change in our 'good' involving beliefs leaves the applicability conditions of 'good' unchanged.

What this shows, and what the OQA makes clear, is that our practice with 'good' is variable. This is why we cannot define 'good', for to do so successfully would require a 'cross section' of such a dynamic practice. This failure does not mean that realism is the only option. My suggestion is that all we can say concerning the meaning of the question 'is this good?' is that it is a question we attempt to answer by means of revisable procedures; I discuss this in more detail below. In addition, all there is to be said about those things to which 'good' is applicable is just that 'good' is applicable to those things which the questions that we attempt to answer by means of these procedures should receive a positive answer. However, as I have suggested, for this approach to be as plausible as moral realism, I have to show how it can capture those features of our practice with the predicate 'good' for which only realists thought only they could account.

One thing that the OQA has shown is that the conclusion that any application of 'good' is *correct by definition* is counter-intuitive. This is something that the reductionist fails to capture. For instance, if we define 'good' as 'that which gives pleasure to x', then if x feels pleasure then the application of 'good' cannot fail to be correct. The realist has the ability to resist this counter intuitive conclusion, as what counts as a correct application of 'good' depends on the instantiation of a property which we may fail to track accurately. The account under consideration can also give the result that the application of 'good' is not correct by definition, as it carries no commitment to the idea that 'good' is definable. However, where it differs from realism is that the inability to define 'good' is not accounted for in terms of a property that we

fail to track. It is simply our practice with 'good' to always leave open the possibility of revision in the procedures that we associate with 'good' which would leave its applicability conditions unchanged, but would force us to change our verdicts on the matter.³⁹⁹ Given this, we can say that the 'fact' we are trying to track with the question 'is this good' will always remain one step ahead of our trying to answer the question, for the possibility will always remain open that we will revise these procedures without changing the meaning of the question. Nevertheless, 'fact' here is not a commitment to the moral facts that the moral realist invokes to account for such a practice with 'good'; it is used in a trivial and non-metaphysical sense.⁴⁰⁰

Another feature of our practice with 'good' is that we think that some revisions of our procedures for applying 'good' would change the conditions under which we can apply it, while others would not. Again, the view I am considering can make room for this. Speakers associate with 'good' certain ways of ascertaining which changes in procedures leave their applicability conditions unchanged. However, these 'ways of ascertaining' are also revisable. On this view, the meaning of questions concerning whether the application condition of 'good' are the same is explained along the same lines as the meaning of questions concerning whether something is good. They are the questions that we attempt to answer by means of our revisable interpretive procedures. As in the case of the question 'is this good?' this account does not entail that the verdicts on sameness of applicability conditions supported by our interpretive procedures are correct by definition. For the possibility always remain open to revise our interpretative

399 Notice this expresses the same idea as conclusion five in §1.

400 Obviously, I will have to show that this is not a mere verbal manoeuvre; I do not discuss this here. For a convincing starting place for such an account, see Skorupski (1999).

procedure which in turn would force us to modify earlier verdicts while leaving the meaning of the question 'does this have the same application conditions?' unchanged.

As it turns out then this view can capture certain features of realism which are highlighted by the OQA, without us having to countenance good as a robust metaphysical property. Of course, on this account we *could* rephrase the question concerning whether 'is this good?' is meaningful as a question concerning whether there is a property whose instantiation conditions determine the applicability of 'good'. Similarly, we *could* rephrase the question whether 'good' is applicable to a state of affairs as the question whether the property, if any, that determines the applicability condition of 'good' is instantiated by the state of affairs. However, this would be a sense of property that does not commit us to a certain metaphysical outlook and hence such a move would not deserve to be called realist on my discussion.⁴⁰¹ Although this raises a number of highly important questions concerning how best to characterize realism, anti-realism, and irrealism, I do not discuss these here. The important claim is that there is a way of accounting for the applicability of the predicate 'good' which does not subscribe to an account of moral realism as characterized in chapters V, VI and VII.

There is a final worry concerning the account under consideration which arises from one view concerning the goal of a theory of reference. Assume that 'good' imposes a condition on states of affairs because it represents the world as being thus and so. What this means, on one

401 Discussing a similar position, Zalabardo writes that our practice with predicates: 'is not characterised by a puritanical rejection of the metaphysical ideology that attends semantic notions on the realist picture. What sets it apart from the realist approach is the order in which semantic and metaphysical notions are construed. The realist starts from a notion of property and its instantiation conditions, and goes on to explain the notion of predicate applicability in terms of a link between predicates and these antecedently available entities... [Whereas the view under consideration] starts by construing the notion of predicate applicability in terms of the procedures that we employ for answering applicability questions. Once the notion of predicate applicability has been construed in this way, he goes on to construe the notion of the property (if any) whose instantiation conditions we are trying to track with our applications of a predicate as a by-product of the semantic notion.' Zalabardo (1996: 149)

theory, is that whether 'good' is applicable depends on whether the state of affairs satisfies this condition. There have to be facts concerning what has to be the case with respect to a state of affairs in order for 'good' to be applicable to it. However, what this view actually commits us to, is two notions of fact. One refers to the facts we assert with our application of 'good', the other refers to facts concerning the applicability conditions of 'good' that make it possible to assert facts with our application of them. The theory of predicate applicability that I reject is one that claims we *must* specify for 'good' what has to be the case with respect to a state of affairs in order for 'good' to be applicable to it. This position does not have to specify for every moral predicate what has to be the case with respect to a state of affairs for a moral predicate to be applicable to that state of affairs. I argue that, in fact, all we need to provide is an account of how conscience inclining us to apply 'good' can give rise to the notion of correct and incorrect applications of 'good'. I suggest that this is *all* there is to be said about the notion of truth and *a fortiori* about the notion of predicate applicability and the property that determines the applicability conditions for 'good'.⁴⁰²

§3. Conscience and Nominalism: My proposal is that we take this account of predicate application as correct, and explain the meaning of the question 'is this good?' by reference to speaker's conscience.

The nature of conscience is complex and my account of it is highly tentative. Roughly, we can understand it in terms of those impressions of what is right or wrong, good or bad which

402 For an interesting discussion concerning how such an account of truth may be developed see for example, Skorupski *forthcoming*, and Horgan & Timmons (2000a).

give rise to inclinations to apply moral predicates. It is *what it is like* for us when we feel inclined to apply 'wrong' to pictures of children being blown to pieces in American air raids; it is *what it is like* for us when we feel inclined to apply 'wrong' when we hear the local BNP party canvassing.

Although I do not discuss the notion of conscience in detail, it is essential that we clarify what it is for something to be an impression of what is right or wrong, good or bad. What impressions of what is right or wrong, good or bad are *not* are some inner mental *objects* - if we fail to grasp this then the position under consideration fails to be distinct from realism.⁴⁰³ To think the contrary is to make something that is essentially subjective, objective. If we think like this then it will be mysterious how we could be infallible as to whether we really do have the impressions that characterise conscience. The point is, we should think of the impressions not as *things* in the sense of mind-independent objects. Such impressions are not like a table, for example, which has its own intrinsic inherent properties independent of what we think. However, it would also be wrong to think that the impressions have no reality – they are not *nothing*.

The important point for the nominalist proposal under consideration is that we are infallible in our apprehension of such impressions that make up our conscience. It is never an open question whether we are really having the impression that something is right or wrong, good or bad. How, then, could our impressions of what is right or wrong, good or bad which characterise our conscience – i.e. those impressions of what is right or wrong, good or bad

⁴⁰³ In that it introduces a universal, albeit a mental one, which is the referent of the general term 'good'. This conception of qualia is also discussed by Zalabardo (2001).

which incline us to apply moral predicates - be infallible? The answer lies in how we just characterised such impressions. We should not think of the apprehension of the impressions as a cognitive enterprise. Our verdicts on their presence are not infallible because we are somehow too close to them for mistakes to arise, or just so brilliant at arriving at our verdicts on their presence. The reason is that facts about impressions of what is right or wrong, good or bad are constituted by the verdicts that we feel inclined to emit.

Notice that this does not make our inclinations arising from our conscience arbitrary. I am not free to choose which impressions I feel inclined to emit. For instance, I am not equally inclined to apply 'good' to child sacrifice as I am to keeping promises, just as I am not inclined to change my judgement that Hitler is 'evil' after a hearty lunch.

Furthermore, notice that this does not render our verdicts on the presence of impressions of what is right or wrong, good or bad immune to revision. I might, for instance, discount my inclinations before lunch because I recognise that I was delirious with hunger. Alternatively, I might recognise that my inclinations to apply 'good' to Enoch Powell, but not Hitler are inconsistent, and hence form a new inclination. The important point is that the revisability of my inclinations to apply 'good' arise because one inclination is cancelled due to my forming another.

Consider how these points about conscience might relate to predicate application. Suppose that when I was young my parents tried to teach me how to apply the predicate 'good'. To do so they show me a picture of a man helping an old lady across the road and said 'good', and they point out that sharing food with my brother is 'good', etc. Suppose that, after a year and thousands of examples, I feel I can recognise the impression of what is good in my

conscience that inclines me to apply 'good'. I think I can recognise in my experience what counts as the application conditions for 'good'.⁴⁰⁴

Given that I want to claim that we can explain the applicability of 'good' purely in terms of conscience, there are two different ways of proceeding. The first way would involve *defining* 'good' in terms of my impression of what is good; hence, whether 'good' is applicable to a situation is determined by my conscience when I consider that situation. Assume that on the 30th of May my conscience inclines me to apply 'good' when I consider giving money to charity, keeping promises and not killing. I wake the next day to find that that my conscience inclines me to apply 'good' to stealing, setting fire to cats and public torture. Now, given that we are infallible in our apprehension of the impression of good, then if I think that I am experiencing the same impressions on the 1st of June and the 30th of May then I am. Moreover, we are forced to accept the counter-intuitive conclusion that my new verdicts concerning stealing etc. are correct.

There then arises a problem for this reductionist proposal, for this is not how we think of our practice with the predicate 'good'. When we say that something is 'good', we want to allow for the possibility that we are mistaken, and that the things in May were good whereas the things in June were not. We want to be able to say this even if *everyone's* conscience inclines them to apply 'good' to the states of affairs that ours does in both May and June. The practice of applying 'good' must respect the fact that it is an open question whether 'good' is applicable to a state of affairs. However, on this first view this is not possible, as whatever I am inclined to

⁴⁰⁴ Notice again that if qualia cannot be thought of as distinct objective things; for then grasping the application conditions would amount to apprehending a universal – we would be realists.

say, goes. Hence, it seems, via the OQA, that the attempt to define 'good' in terms of our conscience is implausible.

However, if we jettison the claim that holds that to explain the application of 'good' in terms of the impression of what is good requires us to define 'good' in terms of that impression, then ethical nominalism looks more plausible. I have shown in the last section an account of the applicability of 'good' that shows that such an account is possible. I argue that we can explain the applicability of 'good' in terms of conscience without introducing a reduction of 'good' to the impression of what is good.

Importantly, such a position would not entail that our application of 'good' must be correct. How would such a position work? Well, we can specify the relationship between moral predicates and impressions of what is right or wrong, good or bad in a similar way to how we characterised the relationship between 'good' and the predicates that describe speakers' 'good'-applying procedures. True, I am infallible about the impressions of good that I now associate with 'good', but the association itself is revisable. I may for instance, conclude that the impression of what is good in my conscience should no longer be, or in fact never should never have been taken as a reliable indicator of the applicability of 'good'.

On such a position then even though my verdicts on the presence of the impression of what is good is right by definition, it is always a significant and meaningful question whether my verdicts on the applicability of 'good' are correct. The presence of the impression in my conscience associated with a state of affairs leaves open the question of the applicability of 'good'.

What this provides then is a rough description of how an ethical nominalist could account for moral predicate applicability without facing the difficulties that invalidated the other accounts. We can explain the application of 'good' purely in terms of the agent's conscience. But his verdicts on the applicability of 'good' would not be treated as correct by definition since the pairings of impressions of what is good with 'good' are always open to revisions that leave the applicability conditions of 'good' unchanged.

This means that we have an account that respects the OQA and as such avoids the problems I discussed concerning reductive accounts. In addition, it does not introduce the problems associated with moral realism: the metaphysical, epistemological and semantic worries which I have discussed above. For instance, such an account would not need to draw any comparison between moral terms and natural kind terms as Richard Boyd does, nor would it have to invoke a Platonic account of good as Robert Adams does.⁴⁰⁵ We avoid the pitfalls of realism and reductionism whilst respecting the lessons learnt from the OQA.

§4. Conclusion.

In this chapter, I attempted to develop an account that I felt was most plausible in light of the OQA. I developed an account of predicates applicability that fitted between reductionism - which construes the notion in terms of speakers' inclinations - and realism, which construes the notion in terms of properties. I attempted to develop a position of predicate applicability that explains the notion in terms of inclinations, while rejecting the idea that this explanation should take the form of a reduction. I presented this suggestion as a vindication of ethical nominalism,

⁴⁰⁵ Boyd (1988) discussed in Chapter VI, and Adams (1999) discussed in Chapter VII.

as it explains predicate reference solely in term of the notion of conscience; by which I mean those impressions of what is right or wrong, good or bad that incline us to apply moral predicates. I suggested we ought to define such impressions in terms of speakers' inclinations to classify things as good, right wrong etc. I concluded that this account is the one we ought to adopt in light of the proceeding chapters.

Bibliography

- Adams, Robert (1979). 'Divine Command Metaethics as Necessary A posteriori', in Paul Helm (ed.), *Divine Commands and Morality*. Oxford: Oxford University Press.
- Adams, Robert (1987). *The Virtue of Faith and Other Essays in Philosophical Theology*. New York: Oxford University Press.
- Adams, Robert (1999). *Finite and Infinite Goods: A Framework for Ethics*. New York: Oxford University Press.
- Ambrose, Alice, and Lazerowitz, Morris (eds.) (1970). *G. E. Moore: Essays in Retrospect*. London: Unwin brothers Ltd.
- Apel, Karl-Otto (1994). 'Can Philosophy of Language Provide the Key to the Foundations of Ethics?', in Michaelis Michael and John O'Leary-Hawthorne (eds.), *Philosophy In Mind*. Netherlands: Kluwer Academic Publishers.
- Appiah, Anthony (1986). *For Truth in Semantics*. Oxford: Basil Blackwell.
- Aristotle, (1976). *Ethics*. London: Penguin.
- Arrington, R (1989). *Rationalism, Realism, and Relativism*. United States of America: Cornell University Press.
- Audi, Robert (1993). 'Ethical Naturalism and the Explanatory Power of Moral Concepts', in Steven Wagner and Richard Warner (eds.), *Naturalism*. USA: Notre Dame Press.
- Ayer, A. (1936). *Language Truth and Logic*. Oxford: Basil Blackwell.
- Baldwin, Thomas (1990) *G. E. Moore*. London: Routledge.
- Baldwin, Thomas (1993). *Editor's Introduction to Principia Ethica*.
- Baldwin, Thomas (ed.) (1993a). *G.E. Moore: Selected Writings*. London: Routledge.
- Ball, S (1991). 'Linguistic Intuitions and Varieties of Ethical Naturalism' *Philosophy and Phenomenological Research*, Vol. LI, No. 1, March, pp. 1-38.
- Berkeley, G. (1996). *Philosophical Works: Including the Works on Vision*, ed. by M. R. Ayers. Guernsey: The Guernsey Press Co./Everyman Publishers.

- Blackburn, S. (1984). 'The Individual Strikes Back', *Synthesis*, 58, pp. 281-302.
- Blackburn, S. (1985). 'Error and the Phenomenology of Value', in T Honderich (ed.), *Morality and Objectivity* (London)
- Blackburn, S. (1993). *Essays in Quasi-Realism*. Oxford: Oxford University Press.
- Blackburn, S. (1993a). 'Realism, Quasi, or Queasy?' in J. Haldane and C. Wright (eds.), *Reality, Representation and Projection*. Oxford: Oxford University Press.
- Boghossian, P. (1989). 'The Rule-Following Considerations', *Mind*, 98: pp. 507-49.
- Boghossian, P. (1991). 'Naturalizing Content', in B. Loewer and G. Rey (eds.), *Meaning in Mind: Fodor and his Critics*. Oxford: Blackwell.
- Bonjour, L. (1985). *The Structure of Empirical Knowledge*. USA: Harvard University Press.
- Bosanquet, B. "Principia Ethica: Critical Note", *Mind*, New Series, Volume 1, Issue 50.
- Boyd, R. (1988). 'How to be a Moral Realist', in S. Darwall, A. Gibbard and P. Railton (eds.) (1997). *Moral Discourse and Practice*. Oxford University Press: Oxford.
- Brandt, R. (1979). *A Theory of Good and the Right*. Oxford: Clarendon Press.
- Brentano, F (1969) *The Origins of Our Knowledge of Right and Wrong*. London: Routledge and Kegan Paul Ltd.
- Brentano, F (1973) *Psychology from an Empirical Standpoint*. L. McAlister, tr. A. Rancurello, D. B. Terrel and L. McAlister (eds.). New York: Humanities Press.
- Brink, D. (1986). 'Externalist Moral Realism', *Southern Journal of Philosophy*, supplement, pp.23-42.
- Brink, D. (1997). 'Moral Motivation', *Ethics*, 108, pp. 4-32.
- Butchvarov, Panayot (1982) "That Simple, Indefinable, Nonnatural Property Good" *The Review of Metaphysics*, September, vol. xxxvi, pp. 51-77.
- Child, W. (1994). *Causality, Interpretation and the Mind*. Oxford: Clarendon Press.
- Copp, D. (2000). 'Milk, Honey, and The Good Life on Moral Twin Earth', *Synthese*, 124: pp. 113-137.

- Crane, T. (1991). 'All the Difference in the world', *Philosophical Quarterly*, 41, pp. 1-25.
- Cuneo, T. (1999). 'An Externalist Solution to the "Moral Problem"', *Philosophy and Phenomenological Research*, LIX: pp. 359-380.
- Dancy, J. (1993). *Moral Reasons*. Oxford: Basil Blackwell.
- Darwall, S. (1983). *Impartial Reasons*. Ithaca: Cornell University Press.
- Darwall, S., Gibbard, A., and Railton, P. (eds.) (1992). 'Toward *Fin de siècle* Ethics: Some Trends', *The Philosophical Review*, 101: pp. 115-189.
- Darwall, S. (1998). *Philosophical Ethics*. Oxford: Westview Press
- Davidson, D. (1984). *Inquiries into Truth and Interpretation*. Oxford: Oxford University Press.
- Dennett, Daniel (1984). *Elbow Room: The Varieties of Free Will Worth Wanting*. Cambridge, MA: MIT Press.
- Dennett, Daniel (1990). *Consciousness Explained*. London: Little, Brown & Co.
- Descartes, R. (1993). *The Essential Descartes*, (ed.) Margaret D. Wilson. USA: Penguin Books.
- Divers, John and Miller, Alexander (1994). 'Best Opinion, Intention-Detecting and Analytic Functionalism', *Philosophical Quarterly*, 44: pp. 239-245.
- Dreier, J. (1990). 'Internalism and Speaker Relativism', *Ethics*, 101: pp. 6-26.
- Ewings, A. C. (1962). *Ethics*. London: English University Press.
- Falk, W. D. (1986). *Ought, Reasons and Morality*. London: Cornell University Press.
- Fisher, A. D. (2000). 'Rejecting Externalism and Internalism as a Response to Michael Smith's "Moral Problem"', *STOA*, 2: pp. 61-65.
- Fodor, J. (1987). *Psychosemantics: the Problem of Meaning in the Philosophy of Mind*. USA: MIT Press.
- Fodor, J. (1990). *A Theory of Content and Other Essays*. Cambridge, Mass: MIT Press.
- Foot, P. (1978). *Virtues and Vices*. Oxford: Blackwell.
- Frankena, W. K. (1939). 'The Naturalistic Fallacy', *Mind*, October: pp. 464-477.

- Frankfurt, Harry (1971). 'Freedom of the Will and the Concept of a Person', *Journal of Philosophy*, 68: pp. 5-20.
- Frege, G. (1891). 'Function and Concept', in Frege (1960) P. Geach and M. Black (eds.), *Translations from the philosophical writings of Gottlob Frege*. Oxford: Oxford University Press.
- Frege, G. (1892). 'On Sense and Reference', in Frege (1960) P. Geach and M. Black (eds.), *Translations from the philosophical writings of Gottlob Frege*. Oxford: Oxford University Press.
- Frege, G. (1960). *Translations from the Philosophical Writings of Gottlob Frege*, (eds.) P.T. Geach and M. Black, Oxford: Oxford University Press.
- Fuller, M. B. (1991). *Truth, Value and Justification*. Newcastle upon Tyne: Athenaeum.
- Gampel, Eric (1996). 'A Defence of the Autonomy of Ethics: Why Value Is Not Like Water', *Canadian Journal of Philosophy*, 26: pp. 191-210.
- Gampel, Eric (1997). 'The Normativity of Meaning', *Philosophical Studies*, 86: pp. 221-242.
- Gampel, Eric (1997a). 'Ethics, Reference, and Natural Kinds', *Philosophical Papers*, XXVI: pp. 147 - 163.
- Geach, P (1956). 'Good and Evil', *Analysis* 17, pp. 33-42.
- Geach, P., and Black, M. (eds.) (1960). *Translations from the philosophical writings of Gottlob Frege*. Oxford: Oxford University Press.
- Gibbard, Alan (1990). *Wise Choices, Apt Feelings*. Cambridge, Mass: Harvard University Press.
- Greco, J (2000). *Putting Skeptics in their Place: The Nature of Skeptical Arguments and their Role in Philosophical Inquiry*, (Cambridge: Cambridge University Press).
- Hale, B. (1993). 'Can There be a Logic of Attitudes?', in J. Haldane and C. Wright (eds.), *Reality, Representation and Projection*. Oxford: Oxford University Press.
- Hare, R. M. (1952). *The Language of Morals*. Oxford: Oxford University Press.
- Harman, G (1977). *The Nature of Morality*. Oxford: Oxford University Press.
- Harman, G. (1977a). 'Ethics and Observation', in S. Darwall, A. Gibbard, and P. Railton (eds.), *Moral Discourse and Practice*. Oxford: Oxford University Press.

- Harman, Gilbert (1994). 'Doubts About Conceptual Analysis', in Michaelis Michael and John O'Leary-Hawthorne (eds.), *Philosophy In Mind*. Netherlands: Kluwer Academic Publishers.
- Hatzimoysis, Anthony (2002). 'Analytic Descriptivism Revisited', *Ratio*, March, pp. 10-22.
- Horgan, T. & Timmons, M. (1992). 'Troubles For New Wave Semantics', *Philosophical papers*, XXI, pp. 221-260.
- Horgan, T. & Timmons, M. (1993). 'Metaphysical Naturalism, Semantic Normativity and Meta-Semantic Irrealism', in E. Villanueva (ed.), *Naturalism and Normativity. Philosophical Issues*, 4.
- Horgan, T. & Timmons, M. (2000). 'Copping Out Of Moral Twin Earth', *Synthese*, 124: pp. 139-152.
- Horgan, T and Timmons, M. (2000a). 'Nondescriptive Cognitivism: Framework for a New Metaethic', *Philosophical Papers*, 29: pp. 121-53.
- Howich, P. (1995). 'Meaning, Use and Truth'. *Mind*, 104: pp. 355-68.
- Howich, P. (1998). *Meaning*. Oxford: Clarendon Press.
- Hume, David (1975). *Enquiries Concerning Human Understanding And Concerning The Principles of Morals*, (ed.) P. H. Nidditch. Oxford: Clarendon Press.
- Hume, David (1978). *A Treatise of Human Nature*, (2nd edtn.) (ed.) L. A. Selby-Bigge and P. H. Nidditch (Oxford: Oxford University Press).
- Jackson, F. (1974). 'Defining The Autonomy of Ethics', *Philosophical Review*, 83: pp. 88-96.
- Jackson, Frank (1987). *Conditionals*. Oxford: Basil Blackwell.
- Jackson, Frank and Petit, Philip (1988). 'Functionalist and Broad Content', *Mind*, 97: pp. 381-400.
- Jackson, Frank (1994). 'Armchair Metaphysics', in Michaelis Michael and John O'Leary-Hawthorne (eds.), *Philosophy In Mind* Netherlands: Kluwer Academic Publishers.
- Jackson, Frank and Petit, Philip. (1995). 'Moral Functionalism and Moral Motivation', *The Philosophical Quarterly*, 45: pp. 20-40.
- Jackson, Frank (1998). *From Metaphysics to Ethics*. United States: Oxford University Press.

- Jackson, Frank (1999). 'Non-Cognitivism, Normativity, Belief', *Ratio*, XII: pp. 420-435.
- Johnson, R. (1999). 'Internal Reasons and the Conditional Fallacy', *The Philosophical Quarterly*, 49: pp. 53-71.
- Johnson, R. (1999a). Commentary on 'Moral Cognitivism and Motivation' <http://www.brown.edu/Departments/Philosophy/bears>
- Kalderon, K. (forthcoming). 'Open Questions and the Manifest Image', *Philosophy and Phenomenological Research*
- Kierkegaard, S. (1985). *Fear and Trembling*. Suffolk: The Chaucer Press.
- Kim, Jaegwon (1993). 'Naturalism and Semantic Normativity' in E. Villanueva (ed.), *Naturalism and Normativity*, (Philosophical Issues, 4). USA: Thomson-Shore.
- Kim, Jaegwon (1997). 'Moral Kinds and Natural Kinds: What's the Difference – for a Naturalist?', *Truth: Philosophical Issues*, 8, pp. 205-215.
- Kitcher, Philip (1992). 'The Naturalists Return', *The Philosophical Review*, 101: pp. 53-114.
- Klemke, E. D. (1969). (ed.) *Studies in the philosophy of G.E. Moore*. Canada: Burns and MacEachern Ltd.
- Kolnai, Aurel (1980). 'The Ghost of the Naturalistic Fallacy', *Philosophy*, 55: pp. 5-16.
- Kripke, S. (1980). *Naming and Necessity*. Oxford: Blackwell.
- Kripke, S. (1982). *Wittgenstein on Rules and Private Language*. Oxford: Oxford University Press.
- Langford, C.H. (1942). 'Moore's Notion of Analysis' in P. A. Schlipp (ed.), *The Philosophy of G. E. Moore*. New York: Tudor.
- Laudan, Larry (1990). 'Normative Naturalism', *Philosophy of Science*, LV11: pp. 44-59.
- Laurence, Margolis and Dawson <http://www.shef.ac.uk/~phil/staff/laurence/papers/MRTE.html>
'Moral Realism and Twin Earth'.
- Levinson, J. (ed.) (1998). *Aesthetics and Ethics*. Cambridge: Cambridge University Press.

- Levy, Paul (1979). *Moore: G. E. Moore and the Cambridge Apostles* London: Weidenfeld and Nicolson.
- Lewis, D. (1969). *Convention: A Philosophical Study*. Cambridge Massachusetts: Harvard University Press.
- Lewis, D. (1972). 'Psychophysical and Theoretical Identifications', *Australasian Journal of Philosophy*, 50: pp. 249 – 58.
- Lewis, D. (1989). 'Dispositional Theories of Value', *Proceedings of the Aristotelian Society*, 63: pp. 129-132.
- Lewis, D. (2001). 'Forget about the 'correspondence theory of truth'', *Analysis*, October, pp. 275-80.
- Lillehammer, H. (1997). 'Smith on Moral Fetishism', *Analysis*, July: pp. 187-195.
- Locke, J. (1975). *An Essay Concerning Human Understanding*, (ed.) P. H. Nidditch. Oxford: Clarendon Press.
- McCulloch, G. (1989). *The Game of the Name*. Clarendon: Oxford University Press.
- McCulloch, G. (1992). 'The Spirit of Twin Earth', *Analysis*, 52, pp. 168-174.
- McCulloch, G. (1994). *Using Sartre*. USA: Routledge.
- McCulloch, G. (1995). *The Mind and Its World*. London: Routledge.
- McCulloch, G. (2003). *The Life of the Mind* London: Routledge.
- McDowell, J. (1978). 'Are Moral Requirements Hypothetical Imperatives?', *Proceedings of the Aristotelian Society*, 52: pp. 13-29.
- McDowell, J. (1979). 'Virtue and Reason', *Monist*, 62: pp. 331-50.
- McDowell, J. (1988). *Mind, Value, and Reality*. London: Harvard University Press.
- McDowell, J. (1996). *Mind and World*. United States of America: Harvard University Press.
- McNaughton, D. (1988). *Moral Vision*. Oxford: Basil Blackwell.
- Mackie, J. (1977). *Ethics: Inventing Right and Wrong*. Harmondsworth: Penguin.

- Mandelbaum, M. (1955). *The Phenomenology of Moral Experience*. Illinois: The Free Press.
- Maund, B. (1995). *Colours*. USA: Cambridge University Press.
- Miller, A. (1996). 'An Objection to Smith's Argument for Internalism', *Analysis*, 56, pp. 169-174.
- Miller, A. (1998). *Philosophy of Language*. London: UCL Press.
- Miller, A. and McFarland, D. (1998a). 'Response-Dependence without Reduction', *Australasian Journal of Philosophy*, 76: pp. 407-425.
- Miller, A. (forthcoming). *An Introduction to Contemporary Meta-Ethics*. Cambridge: Polity Press.
- Moore, G. E. (1903). 'The Refutation of Idealism', *Mind*, XII: pp. 433-453.
- Moore, G. E. (1912). *Ethics*. London: Williams and Norgate.
- Moore, G. E. (1922). *Philosophical Studies*. London: Kegan Paul.
- Moore, G. E. (1922a). 'The Conception of Intrinsic Value', in Moore (1993). Cambridge: Cambridge University Press.
- Moore, G. E. (1927). 'Facts and Propositions', *Proceedings of the Aristotelian Society*. Suppl. VI, pp. 179-89.
- Moore, G. E. (1942). 'A Reply to my Critics', in P. A. Schlipp, (ed.), *The Philosophy of G. E. Moore*. New York: Tudor.
- Moore, G. E. (1953). *Some Main Problem of Philosophy*. London: Macmillan.
- Moore, G. E. (1993). *Principia Ethica: Revised Edition*, (ed.) T. Baldwin (1993). Cambridge: Cambridge University Press.
- Murdoch, Iris (1970). *The Sovereignty of Good*. Reading: Cox and Wyman Ltd.
- Nerlich, G. (1989). *Values and Valuing*. Oxford: Clarendon Press.
- Nowell-Smith, P. H. (1957). *Ethics*. London: Penguin.
- O'Leary-Hawthorne, John (1994). 'Truth-Aptness and Belief', in Michaelis Michael and John O'Leary-Hawthorne (eds.), *Philosophy In Mind*. Netherlands: Kluwer Academic Publishers.

- Phillips, D. (1977). 'In Search of the Moral "Must"', *Philosophical Quarterly*, 27: pp. 140-57.
- Platts, Mark (1981). 'Moral Reality and the End of Desire' in Mark Platts, ed. *Reference, Truth and Reality*. London: Rutledge and Kegan Paul.
- Putnam, H. (1973). 'Meaning and Reference', in A. W. Moore (ed.), (1993), *Meaning and Reference*. Oxford: Oxford University Press.
- Putnam, H. (1975). 'The Meaning of "Meaning"', in *Mind Language, and Reality: Philosophical Papers*. Cambridge: Cambridge University Press.
- Putnam, H. (1981). *Realism, truth, and history*. Cambridge: Cambridge University Press.
- Putnam, H. (1994). 'Sense, Nonsense and the Senses: An Inquiry into the Powers of the Human Mind', *Journal of Philosophy*, vol. 91, pp. 445-517.
- Quine, W. (1951). "Two Dogmas of Empiricism" in *From a Logical Point of View*, New York: Harper & Row, 1963.
- Quine, W. (1960) *Word and Object*. Cambridge, Mass: MIT Press
- Railton, Peter (1986a). 'Moral Realism', *Philosophical Review*, 95: pp. 163-207.
- Railton, Peter (1986b). 'Facts and Values', *Philosophical Topics*, 14: pp. 5-31.
- Railton, Peter (1989). 'Naturalism and Prescriptivity', *Social Philosophy & Policy*, 7: pp. 151-174.
- Railton, Peter (1993a). 'What the Non-Cognitivist Helps Us to See the Naturalist Must Help Us to Explain', in J. Haldane and C. Wright (eds.), *Reality, Representation and Projection*. Oxford: Oxford University Press.
- Railton, Peter (1993b). 'Reply to David Wiggins', in J. Haldane and C. Wright (eds.), *Reality, Representation and Projection*. Oxford: Oxford University Press.
- Railton, Peter (1993c). 'Noncognitivism about Rationality: Benefits, Costs and an Alternative' in Enrique Villanueva (ed.), *Naturalism and Normativity*, *Philosophical Issues*, 4: pp. 36-51.
- Railton, Peter (1995). 'Subject-ive and Objective', *Ratio*, 8: pp. 259-276.
- Railton, Peter (1998a). 'Aesthetic value, moral value, and the ambitions of naturalism' *Aesthetics and Ethics* (eds.) Levinson, Jerrold pp.59-106 USA: Cambridge University Press.

- Railton, Peter (1998b). 'Moral Explanation and Moral Objectivity', *Philosophy and Phenomenological Research*, 58, pp. 175-182.
- Railton, Peter (1999). 'Normative Force and Normative Freedom: Hume and Kant, But not Hume Versus Kant', *Ratio*, XII, pp. 320-353.
- Rawls, John (1971). *A Theory of Justice*. Oxford: Oxford University Press.
- Rohatyn, D. (1982). 'Burying Moore's Naturalistic Fallacy', *Proceedings of The American Catholic Philosophical Association*. Vol. LVI, pp. 173-185.
- Rosati, C. (1995). 'Naturalism, Normativity, and the Open Question Argument', *Nous*, pp. 46-70.
- Rosati, C. (1995a). 'Persons, Perspectives, and Full Information Accounts of the Good', *Ethics*, 105, pp. 296-325.
- Rosen, Gideon (1994). 'Objectivity and Modern Idealism: What Is The Question?', in Michaelis Michael and John O'Leary-Hawthorne (eds.), *Philosophy In Mind*. Netherlands: Kluwer Academic Publishers, pp. 277-319
- Ross, W. D. (1939). *Foundations of Ethics*. Oxford: Clarendon Press.
- Sartre, J. (1956). *Being and Nothingness*. New York: Methuen & Co. Ltd
- Sayre-McCord (1997). "'Good' On Twin Earth" in *Truth: Philosophical Issues*, 8, (ed.) Enrique Villanueva.
- Sayre-McCord (1997a). "Different Kinds of Kind Terms: A Reply to Sosa and Kim", *Philosophical Issues*, 8, (ed.) Enrique Villanueva.
- Scanlon, T. (1998). *What We Owe to Each Other*. Cambridge: Mass.
- Schlipp, P. (1942). (ed.). *The Philosophy of G. E. Moore*. New York: Tudor.
- Shafer-Landau, R. (1998). 'Moral Judgements and Moral Motivation', *The Philosophical Quarterly*, 48, pp. 353-358.
- Sidgwick, Henry (1963). *The Methods of Ethics*, London: Macmillan.
- Simpson, E. (1999). 'Between Internalism and Externalism in Ethics', *Philosophical Quarterly*, Vol. 49, pp. 201-214.

Skorupski, John. (1986). 'Anti-realism: Cognitive Role and Semantic Content', in J. Butterfield (ed.), *Language, Mind and Logic*. Cambridge: Cambridge University Press.

Skorupski, John (1993a). *English-Language Philosophy 1750-1945*. New York: Oxford University Press.

Skorupski, John (1993b). 'Anti-realism, Inference, and the Logical Constants', pp. 133-165, in J. Haldane and C. Wright, (eds.), *Reality, Representation and Projection*. Oxford: Oxford University Press.

Skorupski, John (1997a). 'Meaning, use, verification', in R. Hale and C. Wright (eds.), *The Blackwell Companion to the Philosophy of Language*. Oxford: Blackwell, pp. 29-59.

Skorupski, John (1997b). 'Logical Grammar, Transcendentalism, and Normativity', *Philosophical Topics*, Vol. 25, pp. 189-211.

Skorupski, John (1997c). 'Reasons and Reason' in *Ethical Explorations* (1999a). Oxford: Oxford University Press.

Skorupski, John (1997d). 'Why did Language matter to Analytic Philosophy?' *the Rise of Analytic Philosophy*, (eds.) Glock, Hans-Johann (Oxford: Blackwell).

Skorupski, John (1999). 'Irrealist Cognitivism', *Ratio*, XII 4, pp. 436-459.

Skorupski, John (1999a). *Ethical Explorations*. United States of America: Oxford University Press.

Skorupski, John (1999b). Commentary on 'Moral Cognitivism and Motivation' <http://www.brown.edu/Departments/Philosophy/bears/0107skorhtml>

Skorupski, John (*forthcoming*). 'Propositions about Reasons'.

Smith, M. (1993). 'Objectivity and Moral Realism: On the Significance of the Phenomenology of Moral Experience', in J. Haldane and C. Wright (eds.), *Reality, Representation and Projection*. Oxford: Oxford University Press.

Smith, M. (1994). *The Moral Problem*. Oxford: Basil Blackwell.

Smith, M. and Kennett, J (1994a). 'Philosophy and Commonsense: The Case of Weakness of Will', in Michaelis Michael and John O'Leary-Hawthorne (eds.), *Philosophy In Mind* Netherlands: Kluwer Academic Publishers.

Smith, M. (1996). 'The Argument for Internalism: Reply to Miller', *Analysis*, pp. 175-183.

Smith, M. (2001). 'Some Not-Much-Discussed Problems For Non-Cognitivism in Ethics', *Ratio*, XIV, pp. 94-114.

Sobel, D. (1994). 'Full information accounts of well-being', *Ethics* 104, pp. 784-810.

Sosa, Ernest (1997). 'Water, Drink, and "Moral Kinds"', *Philosophical Issues*, 8, (ed.) Enrique Villanueva.

Stevenson, Charles (1963). *Facts and Values*. USA: Yale University Press.

Stuart, J. (1998). *A Critical Non-Reductionist Realist Response to Kripke's Wittgenstein's Sceptical Paradox*. (Birmingham: University of Birmingham PhD thesis.)

Sturgeon, N. (1984). 'Moral Explanations', in D. Copp and D. Zimmerman (eds.), *Morality, Reason and Truth*. Totowa, NJ: Rowman & Littlefield.

Sturgeon, N. (1986). 'Harman on Moral Explanations of Natural Facts', *Southern Journal of Philosophy*, supplement, pp. 69-78.

Svavarsdóttir, S. (1999). 'Moral Cognitivism and Motivation', *Philosophical Review*, 108, pp. 161-219.

Svavarsdóttir, S. (2001). 'Svavarsdottir Replies to Critics'
<http://www.brown.edu/Departments/Philosophy/bears/0108svav.html>.

Timmons, Mark (1999). *Morality without Foundation: A Defense of Ethical Contextualism*. New York: Oxford University Press.

Tredwell, R. F. (1959). 'On Moore's Analysis of Goodness', in E. D. Klemke, (ed.), (1969) *Studies in the philosophy of G.E. Moore*. Canada: Burns and MacEachern Ltd.

Velleman, J. (1998a). 'Is Motivation Internal to Value?', in C. Fehige and U. Wessels, Arthur Collignon (eds.), *Preferences*. GmbH: Berlin, pp. 88-102.

Villanueva, E. (ed.) (1993). *Naturalism and Normativity*. USA: Ridgeveiw.

Watson, G (1987). "Responsibility and the Limits of Evil: Variations on a Strawsonian Theme" in *Responsibility, Character, and the Emotions* (ed.) Ferdinand Schoeman, Cambridge: Cambridge University Press.

White, A. R. (1958). *G. E. Moore*. Oxford: Blackwell and Mott, Ltd.

White, M. (1952). "The Analytic and the Synthetic: An Untenable Dualism," in Linsky (ed.), *Semantics and Philosophy of Language*. USA: University of Illinois Press.

Wiggins, D. (1984). 'The Sense and Reference of Predicates: A Running repair to Frege's Doctrine and a Plea for the Copula.', *Philosophical Quarterly*, No. 34

Wiggins, D. (1993a). 'Cognitivism, Naturalism, and Normativity: A Reply to Peter Railton', in J. Haldane and C. Wright (eds.), *Reality, Representation and Projection*. Oxford: Oxford University Press.

Wiggins, D. (1993b). 'A Neglected Position?', in J. Haldane and C. Wright (eds.), *Reality, Representation and Projection*. Oxford: Oxford University Press.

Wiggins, D. (1996). 'Objective and Subjective in Ethics, with Two Postscripts on Truth', in Brad Hooker (ed.), *Truth in Ethics*. Cambridge: Basil Blackwell.

Wiggins, D. (1998). *Needs, Values, Truth*. Oxford: Clarendon.

Williams, B. (1985). *Ethics and the Limits of Philosophy*. (Cambridge: Harvard University Press).

Wittgenstein, L. (1922). *Tractatus Logico-Philosophicus*, translated by C. K. Ogden & F. P. Ramsey. London: Routledge and Kegan Paul.

Wittgenstein, L. (1953). *Philosophical Investigations*. Oxford: Basil Blackwell.

Wittgenstein, L. (1974). *Philosophical Grammar*. Oxford: Basil Blackwell.

Wright, C. (1988). 'Moral value, projection, and secondary qualities', *Proceedings of the Aristotelian Society*, 62, pp. 1-26.

Wright, C. (1989). 'Wittgenstein's rule-following considerations and the central project of theoretical linguistics', in A. George (ed.), *Reflections on Chomsky*. Oxford: Basil Blackwell, pp. 233-64.

Wright, C. (1996). 'Truth in Ethics' in Brad Hooker (ed.), *Truth in Ethics*. Cambridge: Basil Blackwell.

Yablo, S. (forthcoming). <http://www.mit.edu/~yablo/rbb.html>

Zalabardo, J. (1996). 'Predicates, Properties and the Goal of a Theory of Reference', *Grazer Philosophische Studien*, 51, pp. 121-161.

- Zalabardo, J. (1997). 'Kripke's Normativity Argument', *Canadian Journal of Philosophy*, 27, pp. 467-88.
- Zalabardo, J. (1998). 'Putting Reference Beyond Belief', *Philosophical Studies*, 91, pp. 221-257
- Zalabardo, J. (2000). 'Realism Detranscendentalized', *European Journal of Philosophy*, 8, pp. 63-88.
- Zalabardo, J. (2001). 'Towards a Nominalist Empiricism' *Proceedings of the Aristotelian Society*, 101, pp. 29-52.
- Zangwill, Nick (2000). 'Against Analytic Moral Functionalism', *Ratio*, 13(3), pp. 275-286.